

UNIVERSIDAD CATÓLICA SANTO TORIBIO DE MOGROVEJO  
FACULTAD DE INGENIERÍA  
ESCUELA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN



APLICACIÓN DE UN SISTEMA DE ALERTA TEMPRANA BASADA EN  
LA MINERÍA DE DATOS PARA IDENTIFICAR PATRONES  
DELICTIVOS EN LA CIUDAD DE CHICLAYO

TESIS PARA OPTAR EL TÍTULO DE  
INGENIERO DE SISTEMAS Y COMPUTACIÓN

LESLY HAYMET JIMÉNEZ BERRÍOS

Chiclayo 17 de diciembre de 2015

**“APLICACIÓN DE UN SISTEMA DE ALERTA TEMPRANA BASADA  
EN LA MINERÍA DE DATOS PARA IDENTIFICAR PATRONES  
DELICTIVOS EN LA CIUDAD DE CHICLAYO”**

**POR:**

**LESLY HAYMET JIMÉNEZ BERRÍOS**

**Presentada a la Facultad de Ingeniería de la  
Universidad Católica Santo Toribio de Mogrovejo  
para optar el título de**

**INGENIERO DE SISTEMAS Y COMPUTACIÓN**

**APROBADA POR EL JURADO INTEGRADO POR**

---

**Mgtr. Juan Antonio Torres Benavides  
PRESIDENTE**

---

**Mgtr. Marlon Eugenio Vilchez Rivas  
SECRETARIO**

---

**Mgtr. María Ysabel Arangurí García  
ASESOR**

### **DEDICATORIA**

A Dios quien es mi guía en cada momento de mi vida, a mis padres y hermana, quienes son el eje de mi vida, los cuales me han brindado su comprensión, apoyo incondicional y el amor a lo largo de toda la vida.

A ellos, muchas gracias de todo corazón.

### **AGRADECIMIENTOS**

A la Mgtr. María Arangurí García, que gracias a sus conocimientos, sus orientaciones, y su motivación han sido fundamentales para haber concluido satisfactoriamente esta investigación. Así mismo a los miembros de la Región Policial Lambayeque por brindarme toda la información requerida para lograr culminar mi trabajo de investigación.

Y a mis compañeros por toda la ayuda brindada.

## ÍNDICE

I.	INTRODUCCIÓN .....	1
II.	MARCO TEÓRICO.....	5
2.1.	Antecedentes Internacionales.....	5
2.2.	Antecedentes Nacionales y Locales.....	11
2.3.	Bases Teóricas .....	13
2.3.1.	Minería de Datos .....	13
2.3.2.	Inteligencia Artificial.....	26
2.3.3.	Criminalística y Criminología.....	32
2.3.4.	Metodología.....	41
III.	MATERIALES Y MÉTODOS.....	46
3.1.	Diseño de la Investigación.....	46
3.2.	Población .....	46
3.3.	Muestra .....	47
3.4.	Muestreo.....	47
3.5.	Variables .....	48
3.6.	Indicadores .....	49
3.7.	Técnicas e Instrumentos de recolección de datos.....	50
3.8.	Técnicas de Procesamiento de datos .....	51
3.9.	Metodología .....	52
IV.	RESULTADOS .....	53
4.1.	Resultado del Pre – Test.....	53
4.2.	FASE 01: CONCEPCIÓN.....	53
4.2.1.	Fase I: Compresión del Negocio.....	57
4.3.	FASE 02: ELABORACIÓN .....	59
4.3.1.	Fase II: Comprensión de los Datos .....	61
4.3.2.	Fase III: Preparación de los datos.....	84
4.3.3.	Fase IV: Modelado.....	88
4.3.4.	Fase V: Evaluación .....	114
4.4.	FASE 03: CONSTRUCCIÓN.....	116
4.5.	FASE 04: TRANSICIÓN .....	118
4.6.	Resultado del Post – Test .....	138
V.	DISCUSIÓN .....	140
VI.	PROPUESTA.....	146
VII.	CONCLUSIONES .....	148
VIII.	REFERENCIAS BIBLIOGRÁFICAS .....	149
IX.	ANEXOS.....	152

## ÍNDICE DE TABLAS

Tabla 01: Comparación entre clasificación supervisada y no supervisada .....	17
Tabla 02: Delito CVCS .....	35
Tabla 03: Delito contra la Familia .....	35
Tabla 04: Delito contra la Libertad .....	35
Tabla 05: Delito contra el Patrimonio.....	36
Tabla 06: Delito contra la Humanidad .....	36

Tabla 07: Artículos de la Constitución para garantizar la Seguridad Ciudadana...	37
Tabla 08: Decretos para garantizar la Seguridad Ciudadana .....	38
Tabla 09: Diseño de contrastación .....	46
Tabla 10: Población de registro de denuncias en 2009. ....	46
Tabla 11: Técnicas para la Recopilación de Datos .....	51
Tabla 12: Resultados obtenidos en el Pre - Test.....	53
Tabla 13: Alcance del Proyecto .....	54
Tabla 14: Requerimientos Funcionales Clave .....	54
Tabla 15: Costos del Proyecto.....	57
Tabla 16: Requerimientos Funcionales Detallados.....	59
Tabla 17: Descripción de los Campos de la Base de Datos .....	61
Tabla 18: Descripción del Dataset seleccionado.....	67
Tabla 19: Decodificación de la variable Edad del implicado .....	68
Tabla 20: Decodificación de la variable Sexo del implicado.....	69
Tabla 21: Decodificación de la variable Estado civil del implicado .....	69
Tabla 22: Decodificación de la variable Condición Anímica del Implicado .....	69
Tabla 23: Decodificación de la variable Nivel de Instrucción del implicado .....	69
Tabla 24: Decodificación de la variable Reincidencia del Implicado .....	70
Tabla 25: Decodificación de la variable Incidencia Horaria.....	70
Tabla 26: Decodificación de la variable Tipo de Arma utilizada durante el desarrollo de un hecho (delito).....	70
Tabla 27: Decodificación de la variable Nombre de Delito .....	70
Tabla 28: Decodificación de la variable Unidad de Intervención.....	72
Tabla 29: Decodificación de la variable Agresión.....	72
Tabla 30: Decodificación de la variable Delito Circunstancia.....	72
Tabla 31: Decodificación de la variable Monto Robado .....	73
Tabla 32: Variable Edad del implicado .....	73
Tabla 33: Variable Sexo del Implicado .....	74
Tabla 34: Variable Estado Civil del Implicado.....	75
Tabla 35: Variable Condición Anímica del Implicado .....	75
Tabla 36: Variable Nivel de Instrucción del implicado .....	76
Tabla 37: Variable Reincidencia del Implicado .....	77
Tabla 38: Variable Incidencia Horaria.....	77
Tabla 39: Variable Tipo de Arma .....	78
Tabla 40: Variable Nombre de Delito .....	79
Tabla 41: Variable Unidad de Intervención .....	79
Tabla 42: Variable Agresión.....	80
Tabla 43: Variable Delito Circunstancia.....	81
Tabla 44: Variable Monto Robado .....	82
Tabla 45: Dataset Definitivo .....	84
Tabla 46: Unificación de los Delito.....	87
Tabla 47: Resultados del Proceso de Clustering .....	90
Tabla 48: Distribución de Reincidencia del Implicado .....	91
Tabla 49: Resultados obtenidos en el Post - Test .....	139
Tabla 50: Resultados de Encuesta Realizada a los Policías.....	141
Tabla 51: Resultados de Encuesta Realizada a los Policías.....	141
Tabla 52: Resultados de Encuesta Realizada a los Policías.....	142
Tabla 53: Tiempos registrados antes y después de la implementación del Sistema	

Tabla 54: Diferencia de las Medias .....	144
--	-----

## ÍNDICE DE FIGURAS

Fig. 01: La relación entre Minería de Datos con otras disciplinas.....	24
Fig. 02: Modelo de Red hacia adelante.....	30
Fig. 03: Modelo de Red hacia atrás .....	30
Fig. 04: Ciclo de Vida del Proyecto .....	42
Fig. 05: Distribución por edad del implicado.....	74
Fig. 06: Distribución por Sexo del implicado.....	74
Fig. 07: Distribución por Estado Civil del implicado .....	75
Fig. 08: Distribución por Condición Anímica del Implicado.....	76
Fig. 09: Distribución por Nivel de Instrucción del Implicado.....	76
Fig. 10: Distribución por Reincidencia del Implicado.....	77
Fig. 11: Distribución por Incidencia Horaria .....	78
Fig. 12: Distribución por Tipo de Arma.....	78
Fig. 13: Distribución por Nombre de Delito .....	79
Fig. 14: Distribución por Unidad de Intervención .....	80
Fig. 15: Distribución por Agresión .....	81
Fig. 16: Distribución por Delito Circunstancia .....	82
Fig. 17: Distribución por Monto Robado .....	83
Fig. 18: Entrenamiento de la Red Neuronal.....	114
Fig. 19: Base de Datos del Sistema.....	117
Fig. 20: Acceso al Sistema .....	118
Fig. 21: Mensaje de bienvenida.....	118
Fig. 22: Mensaje de error .....	119
Fig. 23: Menú Principal.....	119
Fig. 24: Pestaña Sistema y sus Opciones .....	120
Fig. 25: Pestaña Mantenimiento y sus Opciones .....	120
Fig. 26: Mantenimiento Agraviado .....	121
Fig. 27: Mantenimiento Agraviado sin registro de apoderado .....	121
Fig. 28: Mantenimiento Imputado.....	122
Fig. 29: Mantenimiento Apoderado .....	122
Fig. 30: Mantenimiento Policía .....	123
Fig. 31: Estado o Nivel de acceso sobre el Sistema de Alerta Temprana .....	123
Fig. 32: Unidad de Intervención .....	124
Fig. 33: Delito.....	124
Fig. 34: Tipo delito.....	125
Fig. 35: Pestaña Registrar y sus Opciones .....	125
Fig. 36: Denuncia.....	126
Fig. 37: Hecho.....	126
Fig. 38: Pestaña Aplicación y sus Opciones .....	126
Fig. 39: Características Delictivas .....	127
Fig. 40: Resultados de la exportación de las Características Delictivas.....	127
Fig. 41: Perfiles Delictivos.....	128
Fig. 42: Delincuente Transgresor Semiprofesional.....	128
Fig. 43: Delincuente de Violencia Doméstica.....	129
Fig. 44: Ladrones Merodeadores .....	129

Fig. 45: Delincuente Profesional Consumado .....	130
Fig. 46: Semiprofesional del Delito en Contra del Bien Ajeno .....	130
Fig. 47: Transgresores Familiares.....	131
Fig. 48: Lesionador Indolente .....	131
Fig. 49: Ofensor por Impulsividad.....	132
Fig. 50: Agresor de Peligrosidad Extrema .....	132
Fig. 51: Lesionador de Personas .....	133
Fig. 52: Delincuente Profesional Organizado Local .....	133
Fig. 53: Delincuente Semiprofesional en Delitos Especiales .....	134
Fig. 54: Pestaña Consultas y sus Opciones .....	134
Fig. 55: Consultar Imputados .....	135
Fig. 56: Consultar Imputados .....	135
Fig. 57: Consultar Denuncias .....	136
Fig. 58: Consultar Denuncias en Rojo .....	136
Fig. 59: Consultar Denuncias del Día .....	137
Fig. 60: Consultar Denuncias del Mes .....	137
Fig. 61: Pestaña Reportes y sus Opciones.....	137
Fig. 62: Listado de Imputados .....	138
Fig. 63: Listado de Agravados .....	138



## RESUMEN

La investigación denominada “*Aplicación de un Sistema de Alerta Temprana basada en la minería de datos para identificar patrones delictivos en la ciudad de Chiclayo*”, surgió como alternativa de solución a las deficiencias encontradas en el proceso de registro de denuncias y caracterización de Perfiles Delictivos.

El análisis de la información delictiva es vital para los fines de prevención delictiva, por ello se analizó los registros de los delitos cometidos en el pasado involucrando a los autores materiales, el delincuente. Tomando como base la metodología CRISP-DM.

Tuvo como hipótesis que el desarrollo de un modelo de minería de datos como herramienta de apoyo podría contribuir a la caracterización de perfiles delictivos.

Teniendo como objetivo principal Identificar patrones delictivos de la Ciudad de Chiclayo apoyado por la implementación del Sistema de Alerta Temprana basada en la minería de datos, ya que cada vez este proceso es más tedioso pues la información obtenida de cada denuncia respecto al delincuente está desligada, generando pérdidas de tiempo al momento de analizar dicha información, conllevando a crear perfiles delictivos deficientes.

Los resultados de la investigación indicaron que se logró identificar 12 perfiles delictivos diferenciados demostrando que el Sistema de Alerta Temprana es aplicable pues consiguió reducir considerablemente el tiempo dedicado al proceso de registro de denuncias y búsquedas de perfiles delictivos.

En conclusión la implementación del sistema alcanzó su objetivo, permitió a los policías acceder a la información de forma organizada y sistematizada, además logró minimizar la búsqueda de posibles personas sospechosas ante una denuncia.

**PALABRAS CLAVE:** Patrones delictivos, Alerta temprana, Minería de datos, Agrupamiento de datos.

## **ABSTRACT**

The research entitled "Application of Early Warning System based on data mining to identify crime patterns in the city of Chiclayo," emerged as an alternative solution to the deficiencies in the registration of complaints and characterization of criminal profiling.

The analysis of criminal information is vital for the purpose of crime prevention therefore analyzed the records of crimes committed in the past involving the perpetrators of these acts, the offender. Based development methodology CRISP-DM

He was hypothesized that the development of a data-mining model as a support tool could contribute to the characterization of criminal profiling.

Taking as main objective Identify criminal patterns of the Chiclayo city supported by the implementation of the Early Warning System based on data mining, since each time this process is more tedious because the information obtained from each complaint regarding the offender is detached, causing loss of time when analyzing such information, leading to poor criminal profiling.

The research results indicated that it was possible to identify 12 distinct profiling showing that the Early Warning System is applicable since managed to significantly reduce the time devoted to the process of registration of complaints and queries of criminal profiling.

In conclusion, the implementation of the system reached its goal, allowed the police access to information of way organized and systematic, also achieved minimize the search for possible suspects to a complaint.

**KEYWORDS:** Crime patterns, early warning, data mining, data clustering.

## I. INTRODUCCIÓN

En los últimos años el índice delincucional a nivel mundial aumentó, el cual es un problema que en nuestro país cada vez ha generado mayor preocupación social, tanto por su incremento cuantitativo, como por su progresiva peligrosidad cualitativa, ocasionado un fuerte sentimiento de inseguridad en la población.

En América Latina la delincuencia se ubica dentro de un contexto social caracterizado por grupos de personas ubicados dentro de niveles de miseria o pobreza, desempleo, baja escolaridad o analfabetismo, agresiones sexuales y desintegración familiar.

Según informes del INEI, en el año 2010, el 31,3% de la población del país se encontraba en situación de pobreza, es decir, 31 personas de cada 100, tenían un nivel de gasto inferior al costo de la canasta familiar y el 9,8% en condición de extrema pobreza, de los cuales 596 mil son niños menores de 0-4 años y un millón cien mil, entre 5 y 14 años.

También las deficientes condiciones de vida de múltiples familias conlleva a que los niños trabajen, teniendo que el 55,8% de niños/as de 5 a 13 años que trabaja se encuentra en condición de pobreza, además de que el 66,3% de estos niños tienen un hogar constituido por 5 o más miembros, y de los cuales un 31% de niños solo trabaja (INEI 2013).

De cada 14 niños uno trabaja, y casi la mitad de los niños que trabajan en el país lo hacen en condiciones de explotación, dicha situación ha hecho que estos niños crezcan por lo general, con nulas posibilidades de estudio y adquirieran patrones de conducta nocivos, como consumir sustancias tóxicas y la delincuencia (INFANT 2015 citando a INEI 2011).

Además el aumento del índice de violencia familiar se consideró como otro factor ligado a las personas con alto nivel de incidencia delictiva registrándose 22899 casos de violencia en el año 2010.

El incremento delictivo en el Departamento respecto a años anteriores ha sido notorio pues en el año 2008 el total de denuncias a nivel Nacional fue de 151560 y para el año 2009, fue un total de 160848 denuncias registradas, haciendo que el índice delincucional creciera de 9.23% en el año 2008 a 12.16% en el año 2009 respecto a los demás departamentos del país (Direcciones Territoriales–Direcciones especializadas PNP 2010)

Además de los problemas delictivos más graves registrados durante el 2009 por la Región Policial Lambayeque, fueron el delito contra el patrimonio (se incluye hurto, robo, apropiación ilícita entre otros), siendo un total de 3462 denuncias, equivalente al 64.4% de delitos.

Seguido por los delitos contra la vida, el cuerpo y la salud, cuyo registro fue de 900 denuncias, es decir 16.7%, incluyen lesiones y exposición a peligro o abandono de personas en peligro; después estuvieron los delitos contra la libertad, con 852 denuncias, representando un 15.8%, referidos a violación de la libertad personal, violación de la intimidad, violación de domicilio, y finalmente los conformados por el delitos contra la familia con 162 denuncias (3%), principalmente el de violencia familiar

De acuerdo a las entrevistas realizadas en la Región Policial Lambayeque, en la PNP del Norte y los cuadros estadísticos revisados, se pudo evidenciar algunas de las dificultades a las que se enfrentan para erradicar los problemas delictivos, a ello se le suma la inexistencia de herramientas tecnológicas que contribuyan de igual forma a minimizarlos, como los que se mencionan a continuación:

La falta de implementación de equipos tecnológicos y sistemas informáticos en más del 80% de comisarías en Chiclayo generando pérdidas de tiempo de 1 día a 1½ de retraso en el momento de registrar las denuncias en la base de datos, es decir se realiza doble actividad, debido a que dicho proceso en las comisarías se realiza de manera manual y en la Región Policial Lambayeque este proceso es informatizado.

Al ser el proceso de manera manual se incurre en duplicación de datos, en el 2009 del total de registros (5376) se incurrió en un 2% de denuncias con personas reincidentes, es decir existen personas que han sido registradas por más de una denuncia sin embargo no se ha tomado en cuenta para analizar su perfil criminalístico.

En múltiples ocasiones el historial delictivo de un implicado no se ha tomado en cuenta al momento de analizar y establecer un perfil delictivo, es decir no se realizan seguimientos completos de las actividades delictivas de una persona lo cual ha ocasionado que el ámbito de investigación sea limitado y no se consideren variables importantes como el índice de reincidencia, índice de delitos cometidos, etc., generando un proceso aislado de la realidad.

Demoras en la resolución de los casos, dependiendo del caso pueden demorar de 1 a 2 días si la falta es leve y 15 días o prolongarse si es falta grave, así mismo de las 5376 denuncias registradas por la Región Policial Lambayeque en el 2009, solo se resolvieron 4736 denuncias, quedando un 11.9% en espera.

Por ello es que surgió la interrogante ¿De qué manera la Aplicación de un Sistema de Alerta Temprana basada en la minería de datos en la Región Policial Lambayeque podría contribuir a identificar los posibles perfiles delictivos y por ende a minimizar los problemas delictivos de la ciudad de Chiclayo?

De esta manera se obtuvo como hipótesis que el desarrollo de un modelo de minería de datos como herramienta de apoyo podrá contribuir a la caracterización de posibles perfiles delictivos de la ciudad de Chiclayo.

Para ello la investigación tuvo como objetivo principal Identificar patrones delictivos de la Ciudad de Chiclayo apoyado por la implementación del Sistema de Alerta Temprana basada en la minería de datos.

Por consiguiente se consideró como objetivos específicos determinar las necesidades de los agentes involucrados (PNP Norte – Región Policial Lambayeque), realizar un diagnóstico de la situación actual de los problemas delictivos en la ciudad de Chiclayo y finalmente identificar el nivel de satisfacción de los agentes policiales al hacer uso del Sistema de Alerta Temprana.

Desde el punto de vista Económico se justificó debido a que los costos del desarrollo de la investigación fueron accesibles, entonces el desarrollo de un Sistema de Alerta Temprana basado en la minería de datos, es factible, puesto que planteó una nueva forma de ofrecer información, la cual permitió identificar perfiles delictivos de acuerdo a las características que estos poseen, con lo cual se pudo realizar proyecciones y plantear medidas de prevención.

Desde el punto de vista Científico, se justificó la investigación puesto que permitió conocer, diseñar y desarrollar un Sistema de Alerta Temprana basado en minería de datos para identificar patrones delictivos. Además por ser una propuesta innovadora que contribuyó con la PNP Chiclayo, a obtener su información organizada, clasificada y almacenada de esa forma realizar su trabajo con mayor efectividad.

Desde el punto de vista Tecnológico es importante la propuesta del desarrollo del Sistema de Alerta Temprana basado en minería de datos para identificar patrones delictivos de la ciudad de Chiclayo como herramienta de apoyo que permitió lograr la caracterización de los posibles perfiles delictivos (sospechosos), y generar alertas de esta manera se logró la adquisición y utilización de herramientas tecnológicas necesarios para el desarrollo de la investigación.

La presente investigación tiene una importancia social debido a que logró mejorar el acceso a la Información, siendo esta una situación favorable para la Región Policial Lambayeque, la cual logró hacer visible el conocimiento oculto presente en la policía de Chiclayo, desarrollando una cultura intensiva en conocimiento, reduciendo tiempo, mejorando su trabajo y facilitando el acceso a sistemas informáticos colaborativos.

Presenta una justificación sistémica porque el estudio tuvo un enfoque holístico, donde se identificó a todos los agentes involucrados que

aportaron con sus alcances y sugerencias acerca del proceso, todos ellos conformaron el grupo experimental de la investigación, para poder dar un diagnóstico de la situación actual y de acuerdo a ello plantear cambios factibles y deseables acordes con la problemática presentada formulados como propuesta.

Finalmente, dentro de este contexto la aplicación de un sistema de alerta temprana para la explotación de información delictiva permitió obtener patrones de comportamientos que contribuyan a facilitar la generación de estrategias de prevención. Así mismo, buscó estudiar el valor agregado de aplicar la minería de datos en la detección de patrones delictivos a fin de caracterizarlos logrando extraer conclusiones en la prevención de delitos, es decir aplicar la minería de datos para explicar el pasado a través de la información histórica, entender el presente y predecir la información futura.

## II. MARCO TEÓRICO

### 2.1. Antecedentes Internacionales

<b>TÍTULO</b>	<i>"Aplicación de Minería de Datos para la Detección de Anomalías: Un Caso de Estudio"</i>
<b>AUTOR</b>	Cavero Leal, Amia y Sepúlveda Cuevas, Samuel.
<b>AÑO</b>	2009
<b>INSTITUCIÓN</b>	Universidad de La Frontera – Chile
<b>RESUMEN</b>	<p>La Minería de Datos es una de las soluciones de la Inteligencia de Negocios que ayuda a extraer conocimiento a partir de datos que las empresas han generado producto de su negocio. Este conocimiento puede generar aplicaciones de alto valor agregado, si el proceso de Minería de Datos es entendido apropiadamente desde una perspectiva del negocio.</p> <p>En este contexto, se implementó una aplicación de Minería de Datos que detecta anomalías con el fin de identificar posibles abusos en el uso de los principales servicios que otorga la empresa Aguas Araucanía S.A. Los datos analizados corresponden a las facturaciones de servicios de agua potable y alcantarillado de clientes en la ciudad de Lautaro, que fueron almacenados previamente en un Almacén de Datos. Para el análisis se utilizaron algoritmos de Minería de Datos basados en técnicas de clustering, y la metodología CRISP-DM. Como principal resultado, el sistema permitió a la empresa reducir, el tiempo de búsqueda de los posibles fraudes.</p>
<b>Correlación</b>	<p>El presente antecedente utilizó las técnicas clustering (EM y K-medias) para detectar anomalías, es decir determinar perfiles de usuario con comportamientos anómalos, es aquí donde se encuentra su correlación dado que la detección de patrones delictivos también utiliza la técnica K-medias (no supervisada) para establecer grupos diferenciales entre los posibles sospechosos. Otra de las semejanzas del desarrollo de ambas tesis es que utilizaron la misma metodología de desarrollo CRISP_DM, metodología desarrollada específicamente para construir proyectos de Minería de Datos. Sin embargo la presente tesis no solo busca establecer grupos sino determinar los perfiles delictivos (personas con mayor riesgo de cometer un delito) por ello después se aplicó las redes neuronales como técnica supervisada para obtener dichos resultados.</p> <p>Finalmente el resultado se vio reflejado en la construcción de un Sistema, que permitió mantener la información de forma sistematizada y minimizar los</p>

	índices de búsqueda de las posibles personas sospechosas ante una denuncia.
--	---

<b>TÍTULO</b>	<i>"Explotación de Información Aplicada a Inteligencia Criminal en Argentina"</i>
<b>AUTOR</b>	Britos, P., Fernández, E., Merlino, H., Pollo-Cataneo, F., Rodríguez, D., Procopio, C., Rancan, C., García-Martínez, R.
<b>AÑO</b>	2008
<b>INSTITUCIÓN</b>	Universidad de Buenos Aires – Argentina
<b>RESUMEN</b>	<p>El presente trabajo describe un Proyecto de Explotación de información en el ámbito de la información criminal, analizando homicidios dolosos, homicidios culposos en accidentes de tránsito y la población carcelaria Argentina. Para ello se contó con cinco sistemas de información utilizados por la Dirección Nacional de Política Criminal como fuente para la generación de estadísticas oficiales: Encuestas de Victimización, Sistema Nacional de Información Criminal (SNIC), Sistema de Alerta Temprano (SAT), Sistema Nacional de Estadísticas Judiciales (SNEJ), Sistema Nacional de Estadísticas sobre Ejecución de la Pena (SNEEP).</p> <p>El análisis se realizó a través del uso herramientas de explotación de información de distribución libre.</p>
<b>CORRELACIÓN</b>	<p>Este antecedente evaluó 3 tipos de delitos, más frecuentes en Argentina, es aquí la correlación pues al querer determinar los perfiles delictivos también se evalúa un conjunto de características comunes de los sospechosos, además ambas tesis están enfocada a la misma línea de investigación (ámbito delictivo) sin embargo en la presente tesis se consideró otros tipos como son delitos contra el Patrimonio, contra la Familia y contra la Vida el Cuerpo y la Salud, los cuales son más frecuentes según informes de la Región Policial Lambayeque. Otra similitud respecto al antecedente es la utilización de diversas herramientas de explotación de información para llevar a cabo el análisis de los datos, sin embargo el análisis realizado por el antecedente fue gracias a los datos históricos de las bases de datos de los sistemas existentes en la Dirección Nacional de Política Criminales, y respecto a la presente tesis primero se ha tenido que automatizar este proceso, dando lugar a la construcción de la Base de datos, puesto que solo se contaba con informes impresos.</p> <p>Finalmente ambas tesis buscan obtener patrones de comportamientos delictivos que faciliten la generación de estrategias de prevención contra los crímenes y</p>



	delitos, comenzando por los hechos ocurridos hasta llegar a la persona que lo comete.
--	---

<b>TÍTULO</b>	<i>"Caracterización de la Población Carcelaria en Argentina mediante la aplicación de Minería de Datos para la prevención de hechos delictivos"</i>
<b>AUTOR</b>	Gutiérrez Rüegg, Patricio.
<b>AÑO</b>	2008
<b>INSTITUCIÓN</b>	Instituto Tecnológico de Buenos Aires – Universidad Privada. Argentina.
<b>RESUMEN</b>	<p>El análisis de la información criminal que los diferentes entes del Estado Argentino generan es un punto de vital importancia a los fines de prevención del delito. Se cree que no basta solo con analizar los registros sobre los distintos delitos cometidos en el pasado, sino que es necesario poder llevar a cabo estudios que involucren a los autores materiales de dichos hechos. Generalmente se analizan estadísticas relacionadas con los hechos propiamente dichos, pero nunca se hace hincapié en la persona que ejecuta dicho delito, el delincuente.</p> <p>En el presente antecedente se planteó desarrollar y proponer un modelo metodológico basado en minería de datos que ayude a caracterizar a la sociedad carcelaria a fines de obtener patrones de comportamiento y conducta de los delincuentes, basándose en la información criminal, específicamente en los datos de la población carcelaria de Argentina registrados por la Dirección Nacional de Política Criminal dependiente del Ministerio de Justicia y Derechos Humanos de la Nación Argentina.</p>
<b>CORRELACIÓN</b>	<p>Este antecedente permitió caracterizar a la sociedad carcelaria de Argentina, obteniendo patrones de comportamientos y conductas de los delincuentes, a fin de poder extraer conclusiones para la prevención de delitos, es aquí donde encuentra la correlación, pues no solo se busca analizar la información sino prevenir que dichos delitos se realicen, además ambas tesis están ligadas a la misma línea de investigación (ámbito delictivo). Tanto el antecedente como la presente tesis aplicaron técnicas no supervisadas y técnicas supervisadas, primero utilizaron el agrupamiento de datos (clustering), y posteriormente aplicaron algoritmos de clasificación, para identificar reglas de decisión que ayuden a explicar la composición de cada grupo, aunque en las técnicas supervisadas fueron diferentes en cada</p>

	tesis, en el antecedente se utilizó C4.5 y CHAID y en la presente tesis se utilizó redes neuronales sirvieron para obtener patrones de comportamiento comunes entre los imputados (delincuentes).
--	---

<b>TÍTULO</b>	<i>“Aplicación de Minería de Datos para la Exploración y Detección de Patrones Delictivos en Argentina”</i>
<b>AUTOR</b>	Perversi Ignacio
<b>AÑO</b>	2007
<b>INSTITUCIÓN</b>	Instituto Tecnológico de Buenos Aires – Argentina.
<b>RESUMEN</b>	<p>A partir de la crisis de finales de 2001, Argentina se vio afectada por una creciente ola de inseguridad caracterizada por un aumento en los índices delictivos y los niveles de violencia.</p> <p>El análisis de los registros criminales es fundamental en la prevención del delito. Entre otras cosas, porque permite el diseño de políticas y planes de prevención efectivos.</p> <p>Para ello se trabajó en la identificación y detección de patrones de homicidios dolosos cometidos en Argentina durante 2005 en base a información suministrada por la Dirección Nacional de Política Criminal del Ministerio de Justicia y Derechos Humanos de la Nación, organismo encargado de realizar las estadísticas oficiales de criminalidad en Argentina.</p>
<b>CORRELACIÓN</b>	<p>La correlación entre ambas tesis está dada en la identificación de patrones delictivos aplicando técnicas similares como K-medias, el cual marcó el inicio para formar grupos con características similares. Pese a esto, en el antecedente se utilizó técnicas geográfico-visuales: el Mapa del Delito, que permite geo-referenciar los hechos delictivos, obteniendo una visualización geográfica que contempla no sólo la distancia entre hechos, sino también el equipamiento urbano (bancos, comercios, plazas, etc.) y las demarcaciones territoriales (comisarías, barrios, zonas marginales, etc.) medios informáticos utilizados por la Dirección Nacional de Política Criminal del Ministerio de Justicia y Derechos Humanos de la Nación en Argentina, sin embargo en nuestro medio local la Región Policial Lambayeque no cuenta con estos equipos u otros medios sofisticados que permitan monitorear las zonas de mayor problemas delictivo.</p> <p>Sin embargo a pesar de no contar con estas herramientas, ambas tesis han utilizado diversas herramientas de explotación de información para llevar a</p>

	cabo el análisis de los datos, el antecedente utilizó el conocimiento descubierto para reportarlo a las personas interesadas, y respecto a la presente tesis dicho conocimiento se utilizó en la incorporación de un sistema, que permitiera identificar perfiles delictivos.
--	---

<b>TÍTULO</b>	<i>"COPLINK Managing Law Enforcement Data And Knowledge"</i>
<b>AUTOR</b>	Chen Hsinchun, Zeng Daniel, Atabakhsh Homa, Wyzga Wojciech, y Schroeder Jenny
<b>AÑO</b>	2003
<b>INSTITUCIÓN</b>	Universidad de Arizona
<b>RESUMEN</b>	<p>El Proyecto COPLINK creado en el año 1997, financiado por el Instituto Nacional de Justicia y la NSF, desarrollado en el Laboratorio de IA de la Universidad de Arizona en colaboración con el Departamento de Policía de Tucson (TPD) y el Departamento de Policía de Phoenix (PPD).</p> <p>Es un entorno de información y gestión del conocimiento orientada a resolver problemas de compartimiento de información entre agencias policiales y dotar de herramientas necesarias para la recuperación, filtrado de forma inteligente de información pertinente que aún no ha sido depurada.</p> <p>El sistema COPLINK consta de 2 componentes: COPLINK Connect permite a los departamentos de policía compartir datos sin problemas a través de una interfaz fácil de usar que integra diferentes fuentes de datos y COPLINK Detect, encargado de descubrir diversos tipos de asociaciones delictivas existentes en las bases de datos policiales. Dicho proyecto desarrolló 4 tipos de búsqueda disponibles para el usuario: persona, vehículo, incidente, y la ubicación, debido a que estas son las formas de búsqueda que cubren la mayoría de las tareas que realizan los agentes de policía.</p>
<b>CORRELACIÓN</b>	<p>Este proyecto de investigación fue desarrollado para facilitar la distribución de información policial y de gestión del conocimiento, y más que similitudes con la tesis desarrollada, sirvió de fuente de información pues utiliza cuatro tipos de búsqueda de información lo cual fue útil para la tesis desarrollada recientemente.</p> <p>Como correlación entre ambos proyectos es que utilizaron redes neuronales como técnicas supervisadas para predecir un nuevo patrón de comportamiento.</p>

	Finalmente este es un proyecto ampliamente desarrollado, que sirvió como base para el desarrollo del Sistema de Alerta Temprana, dando algunas pautas necesarias a considerar dentro de este campo (ámbito delictivo).
--	--

## 2.2. Antecedentes Nacionales y Locales

<b>TÍTULO</b>	<i>Identificación y predicción de intereses de los alumnos de una facultad de ingeniería aplicando técnicas de agrupamiento.</i>
<b>AUTOR</b>	Mayhuay Tarazona, Liz.
<b>AÑO</b>	2006
<b>INSTITUCIÓN</b>	Universidad de Ingeniería – Lima, Perú.
<b>RESUMEN</b>	<p>Dentro del contexto educativo, y dado la diversidad de ramas que se relacionan con la carrera de sistemas es necesario identificar a que ramas un alumno muestra mayor afinidad. En el proyecto de investigación se planteó la identificación de intereses de los estudiantes de Ingeniería de Sistemas, Computación e informática de la Universidad Nacional de Ingeniería de Lima basándose en el agrupamiento de intereses según las características que muestra un individuo en relación con sus preferencias a uno de los ámbitos de las carreras de sistemas.</p> <p>Utilizó algoritmos de clustering o agrupamiento como C-means y Hard K-means, adicionalmente hizo uso de un test vocacional tradicional con el propósito de identificar grupos según su comportamiento para posteriormente predecir cuáles son esos intereses de los alumnos según sus características innatas.</p>
<b>CORRELACIÓN</b>	<p>Este antecedente utilizó técnicas de clustering para su desarrollo, su correlación está reflejada en la utilización de las mismas técnicas para lograr identificar patrones, como K-medias (técnica no supervisada)</p> <p>Otra similitud es que no solo utilizan técnicas no supervisadas sino que se llegó a un análisis más profundo utilizando técnicas supervisadas, el antecedente utilizó árboles de decisión y la presente tesis para la detección de patrones delictivos se utilizó redes neuronales, ya que es un algoritmo más sofisticado, y que en los últimos años ha alcanzado un desarrollo con mayor interés, lográndose mejorar notablemente.</p>

<b>TÍTULO</b>	<i>Aplicación de una Solución Business Intelligence con Minería de datos para la mejora de toma de decisiones en la Gestión de Matrículas de la Universidad Católica Santo Toribio de Mogrovejo, Lambayeque – Perú.</i>
<b>AUTOR</b>	Barrantes Ruiz, Luis Raúl, Córdova Jiménez, Marvin Genaro.
<b>AÑO</b>	2007
<b>INSTITUCIÓN</b>	Universidad Católica Santo Toribio de Mogrovejo – Chiclayo, Perú.
<b>RESUMEN</b>	<p>Las empresas actualmente tienen la posibilidad de recopilar y almacenar una gran cantidad de datos e información de los clientes. El reto es cómo tomar mejores decisiones en base a la información adquirida. La respuesta a este desafío es Business Intelligence (BI) que ofrece a las empresas la captura de datos originales para luego ser transformados en útiles de tal manera que apoyen a la toma de decisiones.</p> <p>La propuesta planteó el desarrollo de la aplicación Business Intelligence utilizando técnicas de Minería de Datos para mejorar la toma de decisiones analíticas referentes a la Gestión de Matrículas de la Universidad Católica Santo Toribio de Mogrovejo, para poder identificar algunas características no observables que se producirían durante dicha gestión, y que sirva a las autoridades de la mencionada universidad para el soporte de sus tomas de decisiones. A través de este proyecto se buscó lograr la reducción del tiempo de acceso y respuesta en la obtención de la búsqueda de información para generar los reportes de gestión de Matrículas.</p>
<b>CORRELACIÓN</b>	<p>Este antecedente sirvió como fuente para lograr un exhaustivo análisis de información, la limpieza y transformación y la correcta interpretación de los datos procesados, este antecedente desarrolló un pequeño repositorio de datos (data mart) y en la presente tesis se desarrolló una base de datos que contengan todos los datos respecto al proceso de una denuncia el cual permita lograr caracterizar a los implicados en un delito</p> <p>Además su correlación está en que ambos desarrollaron un sistema que sirvió como medio para identificar patrones de comportamiento y dar solución a una situación real de nuestro entorno local.</p>

## **2.3. Bases Teóricas**

### **2.3.1. Minería de Datos**

#### **2.3.1.1. Definición de Minería de Datos**

La Minería de Datos se describe como aquellos métodos avanzados que sirven para explorar y modelizar relaciones en grandes volúmenes de datos y obtener información que se encuentra implícita (Gómez y Suárez 2004).

Otra de las definiciones que se le atribuye a la Minería de Datos es el considerarse un proceso capaz de extraer conocimientos útil y comprensible, previamente desconocido, desde grandes cantidades de datos almacenados en distintos formatos (Hernández, Ramírez, y Ferri 2004 citando a Witten & Frank 2000).

Hoy en día la Minería de Datos, se encarga de integrar numerosas técnicas de análisis de datos y extracción de modelos, capaz de extraer patrones, describir tendencias, predecir comportamientos y aprovechar la información computarizada que nos rodea, permitiendo a los individuos y organizaciones comprender y modelar de manera más eficiente el contexto en que deben actuar y tomar decisiones (Hernández, Ramírez y Ferri 2004).

En conclusión, la gran popularidad alcanzada por la minería de datos en los últimos años como una herramienta útil a permitido su extensión a diversas áreas, a través de la exploración de grandes bases de datos, de manera automática o semiautomática, con la finalidad de encontrar patrones repetitivos capaces de explicar el comportamiento de los datos en un determinado contexto, es decir a partir de la variedad de datos disponibles (materia prima bruta), aplicar una serie de técnicas con la finalidad de lograr obtener información relevante y dar solución a un caso particular.

#### **2.3.1.2. Características de la Minería de Datos**

Vallejos (2006), sostiene que las principales características de la minería de datos son:

- ❖ La característica principal de la minería de datos es que permite descubrir de forma inductiva información y patrones claves que están ocultos en los datos.
- ❖ Explorar los datos que se encuentran en las profundidades de las bases de datos, como almacenes de datos, repositorios que a veces contienen información almacenada durante varios años.

- ❖ Las herramientas de la minería de datos ayudan a extraer el mineral de la información enterrado en archivos corporativos o en registros públicos archivados.
- ❖ La minería de datos es un proceso que invierte la dinámica del método científico, es decir mientras este formula la hipótesis de manera inicial y luego diseña el experimento para coleccionar los datos que confirmen o refuten la hipótesis. La minería de datos, colecciona los datos, para que describan o indiquen por qué son como son y espera que de ellos emerjan la hipótesis (validar la hipótesis).

### 2.3.1.3. Problemas de la Minería de Datos

Acosta (s.f.), define los siguientes problemas:

- ❖ Facilidad con que se puede caer en una falsa interpretación de los datos y la posibilidad de ser engañados por la estadística; ya que la información disponible para la toma de decisiones comúnmente es una serie de promedios y estimadores estadísticos que presentan una generalización de un gran volumen de datos, los cuales al no ser analizados en su totalidad eficientemente, generan conflictos en los resultados.
- ❖ Es una herramienta explorativa y no explicativa, es decir, explora los datos para sugerir hipótesis, sin embargo es incorrecto aceptar dichas hipótesis como explicaciones o relaciones causa-efecto.
- ❖ Existencia de dificultades en el manejo de algunas herramientas informáticas de minería de datos, mostrándose como una desventaja puesto que se necesita de conocimientos acerca de la minería de datos y programación para su utilización.
- ❖ En relación con el tiempo, es muy posible que se deseen hacer inferencias y análisis de datos sobre un período dado, sin embargo, si durante dicho período no se ha registrado el mismo número de variables, éstas no tienen la misma precisión, y/o interpretación, lo cual dificulta la recopilación de las variables.

Es notorio que la minería de datos no es un proceso de hoy sino que se ha ido desarrollando desde hace ya varios años, no obstante su progreso se ha llevado de manera pausada, debido a la presencia de ciertos inconvenientes que han obstaculizado su desarrollo y a la falta de interés en un principio sobre los grandes beneficios que representaba, sin embargo, en la actualidad ha logrado enfocarse a múltiples áreas alcanzando un éxito sin precedentes.



El gran posicionamiento que ha logrado la minería de datos hoy en día se debe a la existencia de diversas técnicas que la conforma, las cuales son capaces de resolver diversos problemas, además del incremento de la existencia de diversas herramientas de software muy útiles que incluyen dichas técnicas las que permiten dar solución a problemas que las empresas de hoy requieren.

#### **2.3.1.4. Clasificación o Tareas de la Minería de datos**

La clasificación es el proceso mediante el cual se agrupa un conjunto de elementos en función de una representación de los mismos en un espacio n-dimensional, el cual ha sido materia de estudio durante muchas décadas y recientemente es considerada una de las tareas fundamentales de la minería de datos (Palma y Marín 2008) **(Ver Anexo 01)**.

Dependiendo de los objetivos del análisis de los datos, presenta dos variantes, la clasificación supervisada y la clasificación no supervisada (Daza 2007).

- ❖ Clasificación supervisada: Llamada también predictiva, es una técnica multivariada que asigna nuevos objetos a grupos previamente definidos, es decir se posee información de a qué clase pertenece cada elemento, y lo que se desea es encontrar un conjunto de “criterios”, probablemente reglas que permitan dado un nuevo elemento, situarlo en un grupo.

Este tipo de clasificación cuenta con un conocimiento a priori, es decir para clasificar un objeto dentro de una categoría o clase se tiene que tener modelos ya clasificados (objetos agrupados con características comunes). Podemos diferenciar dos fases dentro de este tipo de clasificación:

- Primera fase: Presenta un conjunto de entrenamiento o de aprendizaje (para diseño del clasificador) y otro llamado de test o de validación (para clasificación), estos sirven para construir un modelo general para la clasificación.
- Segunda fase: Proceso que se encarga de clasificar los objetos de las que se desconoce la clase a las que pertenecen.

La clasificación supervisada contiene algunas tareas predictivas, entre las más importantes tenemos las siguientes (Hernández, Ramírez y Ferri 2004):

- Clasificación: Tiene por objetivo analizar los datos de entrenamiento y, mediante un método supervisado,

desarrollar una descripción o un modelo para cada clase utilizando las características disponibles en los datos.

- Clasificación suave: Es una extensión de la anterior, en la que se introduce un grado de certeza en la predicción hecha.
  - Estimación de Probabilidades de clasificación: Es una generalización de la clasificación suave.
  - Preferencia o priorización: Determinar a partir de dos o más ejemplos un orden de preferencia.
  - Regresión: Permite formar relaciones entre datos, la diferencia respecto de la clasificación es que la salida es un valor numérico, está quizá es la tarea más sencilla de definir
- ❖ Clasificación no supervisada: Llamada también descriptiva o de descubrimiento del conocimiento, no se tiene mucha información acerca de los grupos, a veces no se sabe siquiera cuantos grupos hay, por lo tanto lo que se busca es encontrar un agrupamiento que reúna en un mismo grupo elementos más “parecidos” y coloque en grupos diferentes a elementos “dispare”.

Este tipo de clasificación permite descubrir tendencias y patrones en los datos actuales sin tener ningún tipo de conocimiento previo acerca de cuáles son los patrones buscados, por tanto el objetivo no es predecir nuevos datos sino describir los existentes.

A diferencia de la supervisada no cuenta con conocimiento a priori, por lo que se debe contar con un área de entrenamiento disponible para la tarea de clasificación.

La clasificación no supervisada suele ser llamada también clustering, presenta un conjunto de objetos con características, de las que no se conoce a qué clase o categoría pertenece, por ello la finalidad es descubrir los grupos de objetos cuyas características afines permitan separar las diferentes clases.

Alternativamente, puede servir como un paso del pre-procesamiento de los datos para otros algoritmos, como por ejemplo, el de clasificaciones en el cual se trabajaría luego sobre los clúster originados.

La clasificación no supervisada contiene algunas tareas, entre las más importantes tenemos las siguientes (Hernández, Ramírez y Ferri 2004):

- Agrupamiento (Clustering): Consiste en agrupar un conjunto de datos sin tener clases predefinidas, basándose en la similitud de los valores de los atributos de los distintos datos. Este tipo de algoritmo se realiza en forma no supervisada ya que no se saben de antemano ni cómo son los grupos ni cuantos hay.

Utilizado en numerosas aplicaciones como reconocimiento de patrones, análisis de datos, procesamiento de imágenes e investigaciones de mercado.

- Correlaciones y factorizaciones: Tiene como objetivo ver si dos o más atributos numéricos están correlacionados linealmente o relacionados de algún otro modo, es decir detecta atributos redundantes o dependencias entre atributos.
- Reglas de asociación: Expresa patrones de comportamiento entre los datos en función de la aparición de valores de dos o más atributos. A diferencia del anterior se caracteriza porque trata con atributos nominales, lo cual es muy frecuente en las bases de datos.
- Dependencias funcionales: A veces se incluye dentro de las reglas de asociación, en realidad, las dependencias funcionales consideran todos los posibles valores, a diferencia de las asociaciones o dependencias de valor.
- Detección de valores e instancias anómalas: Intenta encontrar instancias que no son similares a ninguna otra (o muy pocas) de las existentes, puede ser muy útil para detectar comportamientos anómalos, como fraudes, fallos, intrusos, etc.

**Tabla 01: Comparación entre clasificación supervisada y no supervisada**

Clasificación Supervisada	Clasificación No Supervisada
Llamada Predictiva	Llamada Descriptiva
Cuenta con conocimiento a priori (se tiene información de los grupos o clases)	No cuenta con conocimiento a priori (no se tiene información de los grupos o clases)
Objetivo encontrar un conjunto de criterios, reglas que logren situar un nuevo elemento en un grupo predefinido.	Objetivo agrupar los objetos de acuerdo a la similitud de características separándolos de los objetos dispares.
Cuenta con modelos ya	No se conoce a qué clase

clasificados(objetos agrupados con características comunes)	pertenece y la interpretación de las clases es labor a realizar a posteriori por parte del analista
Se encarga de clasificar los objetos de las que se desconoce la clase a las que pertenecen.	Se encarga de descubrir los grupos de objetos cuyas características afines permitan separar las diferentes clases.
Puede ser realizada posteriormente a la clasificación no supervisada.	Es considera como el primer paso de la clasificación supervisada

### 2.3.1.5. Técnicas de la Minería de Datos

Son implementaciones específicas de los algoritmos que se utilizan para construir un modelo. Estos algoritmos que se utilizan en la resolución de un determinado problema de minería de datos son diferentes uno de otros y cada uno tiene una serie de ventajas e inconvenientes.

Es por ello que al momento de aplicar un determinado algoritmo no solo se toma en cuenta el tipo de problema sino también el tipo de los datos que se está tratando (Pérez y Santín 2006).

Algunas de las técnicas más empleadas en la minería de datos son (Ver anexo 02):

❖ Modelos predictivos: Clasificación.- Modelos basados en aprendizaje supervisado que utilizan un conjunto de datos de entrenamiento para crear el modelo, que posteriormente se utiliza para clasificar individuos desconocidos. Suelen utilizar árboles de decisión y redes neuronales:

- Árboles de decisión: Un árbol de decisión es un modelo de predicción el cual se representa mediante la forma de árbol en donde cada nodo es una decisión, los cuales a su vez generan reglas para la clasificación de un conjunto de datos.

Los árboles de decisión son fáciles de usar, admiten atributos discretos y continuos y tienen como principal ventaja la facilidad de interpretación.

- Redes Neuronales: Son sistemas conexionistas dentro del campo de la inteligencia Artificial, aunque también pueden utilizarse en el campo de la minería de datos, las cuales, dependiendo del tipo de arquitectura neuronal, pueden tener diferentes aplicaciones.

Esta técnica en los últimos años se ha convertido en uno de los instrumentos más frecuente para detectar categorías

comunes en los datos, debido a que son capaces de detectar y aprender complejos patrones, y características de los datos.

Pueden utilizarse para el reconocimiento de patrones, la comprensión de información, agrupamiento, clasificación, visualización, etc.

Este tema será ampliado más adelante por ser de gran importancia para el desarrollo de la presente tesis.

- ❖ Reglas de Asociación: El objetivo de esta técnica es encontrar elementos que implican la presencia de otros elementos dentro una misma transacción, es decir cuenta las ocurrencias de todos los elementos presentes en las transacciones de la base de datos y crear un vector donde cada uno de sus elementos lleva una cuenta de un elemento de la base de datos.
- ❖ Algoritmos Genéticos: Son técnicas de optimización que imitan la evolución de las especies y utilizan procesos tales como combinaciones genéticas, mutación, reproducción y selección, como también proporcionan programas y optimizaciones que pueden ser usadas en la construcción y entrenamiento de otras estructuras como es el caso de las redes neuronales.
- ❖ Agrupamiento o Clustering: Considerada como una de las tareas más frecuentes de la minería de datos, la cual trata de encontrar grupos o clases en un conjunto de individuos.

Esta se divide en Agrupamiento Particional y Jerárquica, el primero forma grupos disjuntos y cubre todo el conjunto de elementos y el segundo es un agrupamiento particional anidado, es decir cada grupo de un nivel se divide en varios en el nivel siguiente, respectivamente.

Dentro de esta división encontramos los siguientes algoritmos:

Agrupamiento Particional: Se considera que los grupos deben ser cohesionados de manera que los elementos de un mismo grupo estén más cercanos entre sí y la distancia entre grupos sea la mayor posible.

Uno de los más extendidos dentro del agrupamiento particional es el de mínimos cuadrados, donde el criterio de cohesión se obtiene como la suma total de la distancia de cada elemento al punto medio (centroide) al grupo al que pertenece.

- Algoritmo K-medias: Es el método que mejor encaja en esta categoría, fue creado por MacQueen en 1967 es el algoritmo

de clustering más conocido y utilizado ya que es muy simple aplicación y eficaz.

El procedimiento K-medias está basado en la utilización de la distancia euclídea, teniendo en cuenta que los centroides se definen como la media aritmética de los centros de cada grupo.

Requiere dos parámetros iniciales: el número de grupos, K, y sus respectivos centroides iniciales, los cuales pueden ser elegidos por el usuario u obtenerse por medio de algún procesamiento previo.

Después de que se dispone de dichos datos, cada elemento es asignado al grupo del centroide más cercano, obteniendo de esta forma la composición inicial de los grupos. Una vez obtenidos los grupos, se recalculan los centroides y se hace una nueva reasignación. El proceso se vuelve a repetir hasta que los centroides no cambian.

- Mapas auto-organizativos o SOM (self-organizing map) de Kohonen: Fueron desarrollados por el científico francés Teuvo Kohonen y son un tipo especial de redes neuronales artificiales de aprendizaje no supervisado exitosamente aplicado como una herramienta de Data Mining.

Los mapas auto-organizativos han tenido una gran variedad de aplicaciones de ingeniería tales como reconocimiento de patrones, análisis de imágenes, monitoreo de procesos, detección de fallas, etc.

La red auto-organizada trata de descubrir rasgos comunes, regularidades, correlaciones o categorías en los datos de entrada, e incorporarlos a su estructura interna de conexiones.

Una red auto-organizada está compuesta por dos capas de neuronas. La capa de entrada (formada por N neuronas, una por cada variable de entrada) recibe y transmite a la capa de salida la información procedente del exterior. La capa de salida (formada por M neuronas) encargada de procesar la información y formar el mapa de rasgos.

Las conexiones entre las dos capas que forman la red son siempre hacia delante, es decir, la información se propaga desde la capa de entrada hacia la capa de salida. Además cada neurona de entrada  $i$  está conectada con cada una de las neuronas de salida  $j$  mediante un peso  $w_{ji}$ .

Agrupamiento Jerárquico: Es una sucesión de particiones “anidadas” donde cada grupo de elementos perteneciente a una determinada partición está totalmente incluido en algún grupo de la partición siguiente (Palma y Marín 2008).

Los algoritmos de agrupamiento jerárquico son de tipo aglomerativo, de forma que partiendo de una partición en la que cada elemento forma un grupo se van obteniendo nuevas particiones uniendo grupos entre sí.

Un enfoque para el desarrollo de estos algoritmos es el basado en grafos, que considera que cada elemento es un vértice de un grafo y se van generando particiones, conectando los vértices de menor distancia

Existen dos formas de agrupamiento:

- Agrupamiento de enlace simple: Los grupos se obtienen buscando las componentes conexas de grafo y se termina cuando todos los vértices están conectados.
- Agrupamiento de enlace completo: Los grupos se obtienen buscando los subgrafos completamente conectados, es decir tendremos un grupo de dos vértices si hemos colocado una arista, el algoritmo termina cuando hemos incluido todos los vértices en un grupo.

#### 2.3.1.6. Algoritmo K-medias (K-means)

K-medias es uno de los algoritmos de aprendizaje no supervisado más simples y más conocido dentro de los algoritmos de agrupamiento, sigue una forma fácil y simple para dividir una base de datos dada en  $k$  grupos (fijados a priori). Los datos se asignan a cada centro según la menor distancia, normalmente usando la distancia euclídea (Luis Vázquez, s.f.)

Básicamente este algoritmo busca formar clúster (grupos) los cuales serán representados por  $K$  objetos. Cada uno de estos  $K$  objetos es el valor medio de los objetos que pertenecen a dicho grupo.

Este algoritmo de agrupamiento comienza particionando los datos en  $k$  grupos. Luego calcula el centroide de cada partición como el punto medio del clúster y asigna cada dato al clúster cuyo centroide sea el más próximo.

- ❖ **Objetivo del algoritmo:** El objetivo que se intenta alcanzar es la mayor diferencia entre grupos establecidos. Para lo cual

después de haberse establecido los centroides estos son recalculados para los nuevos grupos y el algoritmo se repite hasta que no haya más datos que cambien de grupo de una iteración a otra.

Se intenta encontrar entre todas las particiones de  $X$  en  $k$  conjuntos  $\{X_i; i = 1, 2, \dots, k\}$  aquella que minimiza el criterio de agrupamiento.

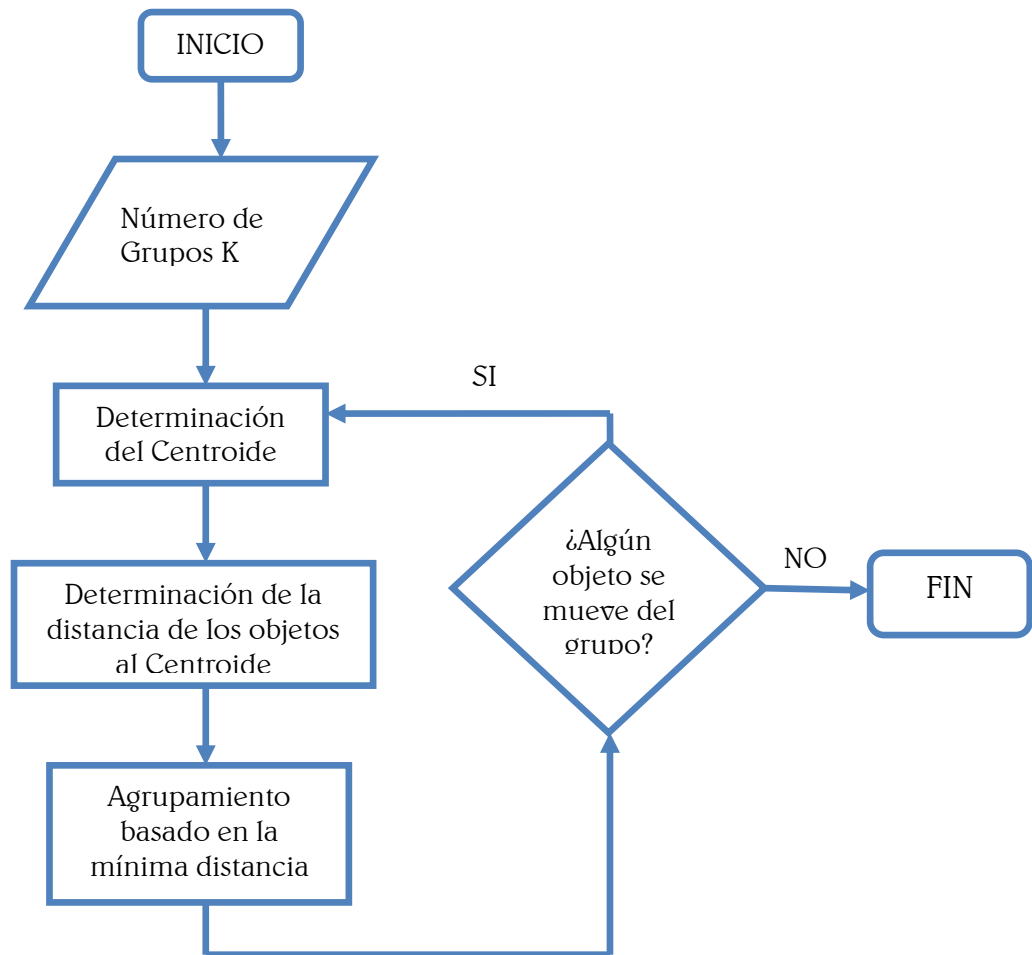
❖ **Algoritmo K-means se implementa en:**

**1. Inicialización:** Inicialmente se seleccionan  $K$  objetos del conjunto de entrada. Estos  $K$  Objetos serán los centroides iniciales de los  $K$ -grupos.

**2. Asignación y actualización de centros:** Se calculan los centroides de los clúster de cada partición (punto medio del clúster) y se asignan cada objeto al clúster cuyo centroide sea más cercano, es decir se calculan las distancias de los objetos (datos) a cada uno de los centroides y se asignan a aquellos grupos cuya distancia es mínima con respecto a todos los centroides.

**3. Convergencia:** En el paso previo algunos patrones pueden cambiar de agrupamiento y en consecuencia los centros de éstos. Es por ello que se recalculan hasta cuando no haya más reasignaciones, es decir hasta que los centroides no se mueven más y termina el agrupamiento.





#### ❖ Ventajas del Algoritmo K-means

- El proceso de agrupamiento k-medias es el algoritmo de clustering (agrupamiento) más conocido y utilizado ya que es de muy simple aplicación y eficaz.
- Escalabilidad: No genera inconvenientes al correrlo con pocos datos aunque también funciona para una gran cantidad de procesamiento de datos.
- Capacidad de manejar diferentes tipos de atributos: numéricos (lo más común), binarios, nominales, ordinales, etc.
- La representación mediante centroides tiene la ventaja de que tiene un significado gráfico y estadístico inmediato.

#### ❖ Desventajas del Algoritmo K-means

- El problema del empleo de estos esquemas es que fallan cuando los puntos de un grupo están muy cerca del centroide de otro grupo.
- Muestra problemas cuando los grupos tienen diferentes tamaños y formas.

- Necesidad de inicializar el número de prototipos al principio de la ejecución. Esto perjudica la eficacia del algoritmo ya que en la práctica, no se conoce a priori el número de clúster final.

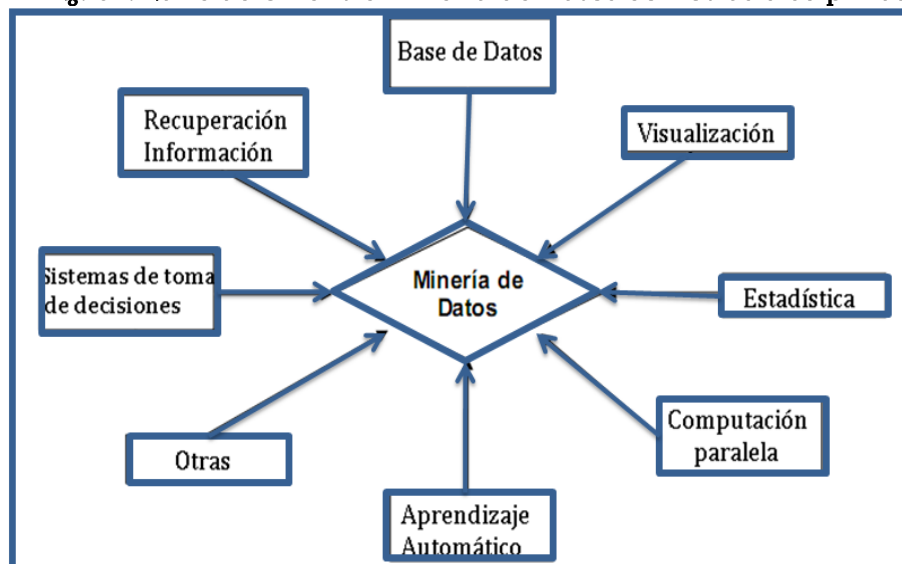
### 2.3.1.7. Minería de Datos relacionada con otras disciplinas

La Minería de Datos es un campo multidisciplinar que se ha desarrollado en paralelo o como prolongación de otras tecnologías. Por ello, la investigación y los avances en la Minería de Datos se nutren de los que se producen en estas áreas relacionadas (Hernández, Ramírez y Ferri 2004).

La introducción de la minería de datos a otras áreas ha logrado una mayor utilidad, a pesar de ello día a día los avances y descubrimientos van ampliándose, haciendo que la minería de datos logre una mayor interacción con diversas áreas importantes en el mundo de hoy.

Dentro las áreas más resaltantes encontramos las siguientes:

Fig. 01: La relación entre Minería de Datos con otras disciplinas



Fuente: Hernández, Ramírez y Ferri (2004).

**Base de Datos:** Incluye conceptos como almacenes de datos y procesamiento analítico en línea (OLAP), los cuales ayudan a extraer conocimiento novedoso y comprensible.

**Recuperación de Información:** Busca obtener información desde datos textuales, basando su desarrollo histórico en el uso de bibliotecas (actualmente bibliotecas digitales) y en la búsqueda por Internet.

**Visualización de datos:** Permite al usuario descubrir, intuir o entender patrones que serían difíciles de ver a través de descripciones textuales o matemáticas de los resultados.

**Sistemas para la toma de decisiones:** Son herramientas y sistemas informatizados que tienen como objetivo proporcionar información necesaria para realizar decisiones efectivas en el ámbito empresarial, es decir ayudan a los directivos en la resolución de problemas y toma de decisiones.

**Computación paralela y distribuida:** Incluyen tecnologías de procesamiento paralelo, distribuido o computación en grid, lo cual permite repartir entre diferentes procesadores o computadores las tareas más complejas de minería de datos.

**Estadística:** Considerada como la madre de la minería de datos, ya que ha proporcionado muchos algoritmos, técnicas utilizadas la minería de datos.

**Aprendizaje Automático:** Área de la Inteligencia artificial que se ocupa del desarrollo de algoritmos (programas) capaces de aprender, y junto a la estadística constituyen el corazón del análisis inteligente de los datos.

**Otras disciplinas:** La relación de la minería con el resto de disciplinas depende de los tipos de datos a ser minados, como el uso de técnicas de análisis de imágenes, procesamiento de señales.

#### **2.3.1.8. Aplicaciones de la Minería de Datos**

Las aplicaciones respecto a la minería de datos están incorporadas en múltiples organizaciones como empresas, gobiernos, hospitales, universidades y otras instituciones interesadas en explorar sus bases de datos, aunque este desarrollo es más notorio a nivel internacional (Vallejos 2006).

❖ En el Gobierno:

El FBI analiza las bases de datos comerciales para detectar terroristas. A principios de julio de 2002, el director del Federal Bureau of Investigation (FBI), John Aschcroft, anunció que el Departamento de Justicia comenzaría a introducirse en la amplia cantidad de datos comerciales referentes a los hábitos y preferencias de compra de los consumidores, con el fin de descubrir potenciales terroristas antes de que ejecuten una acción.

Los expertos aseguraron que, con dicha información, el FBI unirá todas las bases de datos mediante el número de Seguridad Social y permitirá saber si una persona fuma, qué talla y tipo de ropa usa, su

registro de arrestos, su salario, su altura y peso, contribuciones a la Iglesia, grupos políticos u organizaciones no gubernamentales, los libros que lee, los productos de supermercado que compra, si tomó clases de vuelo o si tiene cuentas de banco abiertas, entre otros.

Otra de las aplicaciones realizadas a nivel mundial respecto a los problemas delictivos, es el Proyecto COPLINK creado en el año 1997, financiado por el Instituto Nacional de Justicia, desarrollado en la Universidad de Arizona en colaboración con el Departamento de Policía de Tucson (TPD) y el Departamento de Policía de Phoenix (PPD).

El proyecto COPLINK estuvo orientado a resolver problemas como el compartimiento de información entre agencias policiales, a través de una interfaz fácil de usar que se integra desde diferentes fuentes de datos, así mismo orientado a descubrir diversos tipos de asociaciones delictivas que existen en bases de datos policiales. (Chen, et al. 2003).

### **2.3.2. Inteligencia Artificial**

#### **2.3.2.1. Definición de Inteligencia Artificial - IA**

Existen múltiples definiciones para describir la Inteligencia Artificial, entre ellas tenemos:

Inteligencia Artificial es conocida como el desarrollo de métodos y algoritmos que permiten comportarse a las computadoras de modo inteligente (Gutiérrez 2000).

Otra de las definiciones, es la mencionada por Oyarzo (s.f.) quien sostiene que es el campo de la ciencia que trata de realizar a través de máquinas computacionales tareas que pueden ser realizadas por el hombre, como percibir, razonar y actuar.

#### **2.3.2.2. Definición de Redes Neuronales Artificiales -RNA**

Hilera y Martínez (2000) sostienen que existen numerosas formas de definir lo que son las redes neuronales o computación neuronal.

Se puede definir como una nueva forma de computación inspirada en modelos biológicos.

Otra definición es que son un modelo matemático compuesto por un gran número de elementos procesales organizados en niveles.

Pérez (2007) sostiene que una red neuronal es un conjunto de elementos de procesamiento de la información interconectada, capaz de aprender con la información que se les alimenta.

Finalmente una definición más completa es que son redes interconectadas masivamente de elementos simples las cuales intentan interactuar con objetos del mundo real del mismo modo que el sistema nervioso biológico, no obstante no deja de ser una herramienta creada por el hombre, con cierto grado de inteligencia pero con limitaciones que sólo poseen un parecido superficial con sus contrapartidas biológicas.

### 2.3.2.3. Elementos de una Red Neuronal - RN

Pérez (2007) sostiene que una red neuronal está formada por unidades de procesamiento que reciben el nombre de neuronas o nodos. Estos nodos están organizados en grupos que se llaman “capas”. Generalmente existen 3 tipos de capas: una capa de entrada, una o varias capas ocultas y una capa de salida, las cuales están conectadas.

- ❖ La capa de entrada, mediante la cual se presentan los datos a la red, formada por nodos de entrada que reciben información del exterior.
- ❖ Las capas ocultas se encargan de procesar la información y se interponen entre la capa de entrada y la de salida, están no tienen conexión al exterior.
- ❖ La capa de salida representa la respuesta de la red una entrada dada siendo esta información transferida al exterior.

Cada neurona  $i$ -ésima está caracterizada por un valor numérico denominado estado de activación  $a_i(t)$ ; asociado a cada unidad, existe una función de salida,  $f_i$ , que transforma el estado actual de activación en una señal de salida,  $y_i$ . Dicha señal es enviada a través de los canales de comunicación unidireccionales a otras unidades de la red; en estos canales la señal se modifica de acuerdo con la sinapsis (el peso  $w_{ji}$ ) asociada a cada uno de ellos según una determinada regla. Las señales moduladas que han llegado a la unidad  $j$ -ésima se combinan entre ellas, generando así la entrada total,  $Net_j$  (Hilera y Martínez 2000).

$$Net_j = \sum_i y_i w_{ji}$$

Una *función de activación*,  $F$ , determina el nuevo estado de activación de la neurona, teniendo en cuenta la entrada total calculada y el anterior estado de activación, estas actualizaciones de los estados de las unidades pueden ser de modo asíncrono (neuronas evalúan su estado continuamente, conforme les llega la

información y lo hacen de forma independiente) y síncrono (la información también llega de forma continua, pero los cambios son realizados de forma simultánea) (Hilera y Martínez 2000).

❖ **Unidades de proceso** (Hilera y Martínez 2000).

– Estado de activación:

El procesamiento que realiza la red se ve como la evolución de un patrón de activación en el conjunto de unidades que lo componen a través del tiempo.

Todas las neuronas que componen una red tienen cierto estado, reposo y excitado, las cuales se les asigna un valor (continuo o discreto).

Si son discretos, suelen tomar valores binarios, el cual un estado activo se indica por un 1, caracterizado por la emisión de un impulso por parte de la neurona, y un estado pasivo se indica por un 0 (estado de reposo). Otros modelos consideran un conjunto continuo de estados de activación, asignando valores entre  $[0,1]$  o  $[-1,1]$ , siguiendo una función sigmoideal.

Finalmente, es necesario saber qué criterios siguen las neuronas para alcanzar tales estados de activación:

- a) Se debe tener idea del mecanismo de interacción entre las neuronas, es decir el estado de activación está fuertemente influenciado por tales interacciones
- b) La señal que envía cada una de las neuronas a sus vecinas dependerá de su propio estado de activación.

– Función de salida o de transferencia:

Entre las neuronas que forman una red neural artificial existe un conjunto de conexiones que unen unas a otras.

Cada neurona transmite señales a aquellas que están conectadas con su salida. Además cada neurona tiene asociada una función de salida que transforma su estado actual de activación en una señal de salida.

Existen funciones de transferencia típicas que determinan distintos tipos de neuronas:

- a) Función escalón o umbral: Únicamente se utiliza cuando las salidas de la red son binarias. La salida se activa cuando el estado de activación es mayor o igual que cierto valor origen.

- b) Función lineal o identidad: Es equivalente a no usar la función de salida, pues la salida es igual al nivel de activación, se usa muy poco.
  - c) Función mixta: Es la más apropiada cuando se quiere como salida información analógica, en la cual si la suma de señales de entrada es menor que un límite inferior, la activación es 0 o -1 y si dicha suma es mayor o igual que el límite superior, entonces la activación es 1.
  - d) Función sigmoideal: También se considera entre las más apropiadas cuando se quiere como salida información analógica.
- Conexiones entre neuronas:  
Las conexiones que unen a las neuronas de un RNA tienen asociado un peso, que es el que hace que la red adquiera conocimiento.

Una neurona recibe un conjunto de señales que le dan información del estado de activación de todas las neuronas con las que se encuentra conectada.

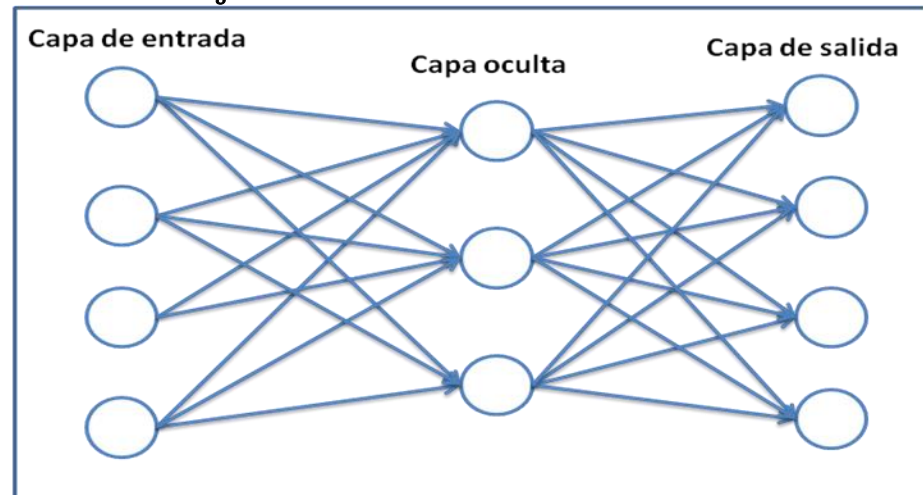
- Función o regla de activación:  
Permite combinar las entradas con el estado actual de la neurona para producir un nuevo estado de activación, es decir se produce un nuevo estado de activación en una neurona a partir del estado que existía y la combinación de las entradas con los pesos de las conexiones.  
Existen diversas funciones de activación:
- a) Función escalón o umbral: Si el conjunto de los estados de activación está entre  $[0,1]$  o  $[-1,1]$ .
  - b) Función lineal o identidad: El conjunto de estados puede contener cualquier número real; el estado de activación coincide con la entrada total que ha llegado a la unidad.
  - c) Función mixta: El estado de activación de la unidad está obligado a permanecer dentro de un intervalo de valores reales prefijados.
  - d) Función sigmoideal: Es una función continua, por tanto el espacio de los estados de activación es un intervalo del eje real.
- Regla de aprendizaje:  
Las RNA pueden considerar que el conocimiento se encuentra representado en los pesos de las conexiones entre neuronas. Todo proceso de aprendizaje implica cierto número de cambios en estas conexiones. En realidad, se puede decirse que el aprende modificando los valores de los pesos de la red.

#### 2.3.2.4. Estructura de una Red Neuronal

Existen dos tipos de estructura de redes neuronales (Pérez 2007):

- ❖ La red alimentada hacia adelante: Es la más habitual, ya que las conexiones entre neuronas se establecen en un único sentido, en el siguiente orden: capa de entrada, capa (s) oculta (s) y la capa de salida.

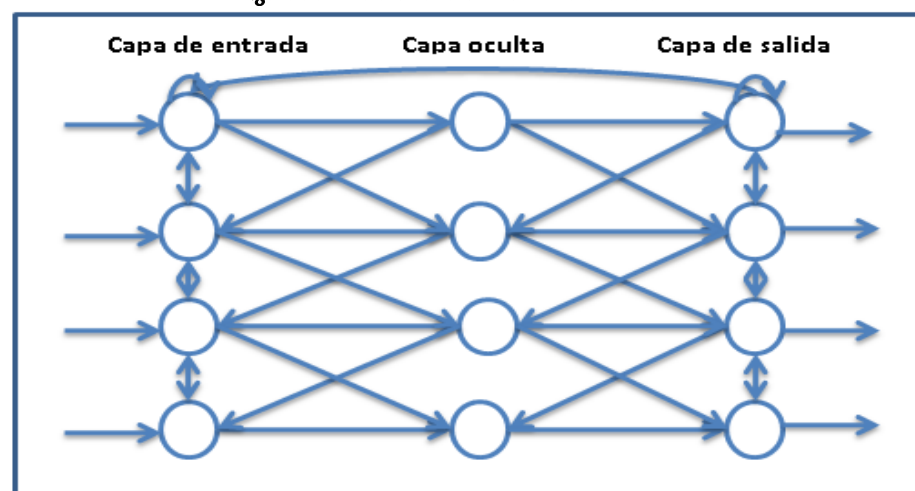
Fig. 02: Modelo de Red hacia adelante



Fuente: Pérez (2007).

- ❖ Las redes retroalimentadas: Tienen conexiones hacia atrás, es decir, de nodos de una capa a elementos de proceso de capas anteriores, logrando poseer conexiones, tanto entre neuronas de una misma capa, como de un nodo a sí mismo.

Fig. 03: Modelo de Red hacia atrás



Fuente: Pérez (2007).

#### 2.3.2.5. Ventajas de las Redes Neuronales



Hilera y Martínez (2000) afirman que las redes neuronales presentan múltiples ventajas:

- ❖ **Aprendizaje adaptativo:** Característica más importante de las redes neuronales, estas aprenden a llevar a cabo ciertas tareas mediante un entrenamiento con ejemplo ilustrativos.

Son adaptables debido a la capacidad de autoajustarse de los elementos procesales (neuronas) que componen el sistema.

- ❖ **Autoorganización:** Usan su capacidad de aprendizaje adaptativo para autoorganizar la información que reciben durante el aprendizaje y/p la operación.

Esta consiste en la modificación de la red neuronal completa para llevar a cabo un objetivo específico.

- ❖ **Tolerancia a fallos:** Las redes neuronales permiten que si se produce un fallo en un pequeño número de neuronas, aunque el sistema se ve influenciado, no sufre ninguna caída repentina.

- ❖ **Operación en tiempo real:** las redes neuronales permiten realizar grandes procesos con datos de forma muy rápida debido a su implementación paralela.

- ❖ **Fácil inserción de la tecnología existente:** Debido a que una red puede ser rápidamente entrenada, comprobada, verificada y trasladada a una implementación hardware de bajo costo, es fácil insertar redes neuronales para aplicaciones específicas dentro de sistemas existentes.

#### **2.3.2.6. Funcionamiento de una Red Neuronal**

Para la creación y aplicación de una red neuronal se debe distinguir los siguientes pasos (Pérez 2007):

- ❖ **Conceptualización del modelo para el estudio del problema concreto:** Aquí se debe señalar las entradas, las salidas y la información de que se dispone.
- ❖ **Adecuación de la información e que se dispone a la estructura de la red a crear:** Se constituyen los patrones de aprendizaje, parte de la información que va hacer utilizada para el entrenamiento o aprendizaje de la red y los patrones de validación, parte de la información que va hacer utilizada como validación de la red.

- ❖ Fase de aprendizaje: Se van presentando los patrones adecuados a la red y esta va proporcionando una salida, este proceso se repite un cierto número de etapas, estas salidas se comparan con las salidas esperadas y los diversos algoritmos de aprendizaje intentan minimizar el error que hay entre la salida proporcionada por la red y la salida esperada
- ❖ Fase de validación: Se presentan a la red entrenada el conjunto de patrones de validación, y se ve el error cometido por la red en este conjunto, este error es una medida de los beneficios de la red.
- ❖ Fase de generalización: Si se ha conseguido una red adecuada se procede a utilizar la red como modelo predictor, aportándole una nueva entrada, la red la procesará y dará una salida.

### **2.3.3. Criminalística y Criminología**

El enorme progreso tecnológico experimentado por nuestro siglo ha acarreado, el nacimiento y el desarrollo de ciencias cuyos orígenes no se remontan más allá del siglo pasado.

En este último caso se encuentran la Criminalística y la Criminología que sirven de auxiliares de la justicia para la determinación de las posibles causales de los delitos ocurridos (Martínez 2009)

#### **2.3.3.1. Definición de Criminalística**

La criminalística es una disciplina que aplica fundamentalmente los conocimientos, métodos y técnicas de investigación, de las ciencias naturales, en el examen de material sensible, significativo y relacionado con un presunto hecho delictivo, con el fin de determinar, en auxilio de los órganos de administrar justicia, su existencia, reconstruirlo o bien señalar y precisar la intervención de uno o varios sujetos en el mismo (Martínez 2009 citando a Montiel 1997).

La Criminalística es una ciencia penal auxiliar, que mediante la aplicación de sus conocimientos, metodología y tecnología descubre y verifica científicamente la existencia de un hecho presuntamente delictuoso y al o los responsables aportando las pruebas a los órganos que procuran y administran justicia (Martínez 2009 citando a Montiel 1997).

En conclusión podemos decir que la criminalística es la ciencia forense que estudia el delito como conducta humana y social, la cual investiga las causas de la delincuencia, de la prevención del delito y del tratamiento del delincuente.

### **2.3.3.2. Definición de Criminología**

Martínez (2009) sostiene que es la ciencia social que estudia la naturaleza, extensión y causas del delito; características de los delincuentes y de las organizaciones delictivas; problemas de detención y castigo de los delincuentes; operatividad de las prisiones y de otras instituciones carcelarias; rehabilitación de los convictos, tanto dentro como fuera de prisión, y la prevención del delito.

El presente trabajo de tesis, aunque evidentemente se trata de una investigación dentro del marco de la Ingeniería de Sistemas y, en particular, en el área de Minería de Datos e Inteligencia Artificial, puede también ser ubicado dentro de esta ciencia auxiliar de la justicia, ya que se busca definir perfiles delictivos que determinen el comportamiento de la actividad delictiva que se presenta dentro del área de investigación.

### **2.3.3.3. Análisis delictivo**

En la actualidad, no existe una definición universalmente aceptada de lo que es Análisis Delictivo. En algunos departamentos de policía es considerado como el estudio de reportes policiacos y la extracción de información que permita la captura de delincuentes, en particular, asesinos seriales (Martínez 2009 citando a Osborne y Wernicke 2006)

En otras agencias, el análisis delictivo consiste en extraer datos estadísticos de las bases de hechos delictivos que ocurren dentro un área y dividirlos en familias delictivas y épocas del año.

Cualquiera que sea la definición de Análisis Delictivo, éste tiene por objetivo, encontrar información importante dentro de los datos contenidos en cada uno de los hechos delictivos y diseminarla entre los oficiales e investigadores para ayudarlos en la captura de posibles delincuentes, así como frenar la actividad delictiva (Martínez 2009 citando a Osborne y Wernicke 2006).

La definición formal de Análisis Delictivo empleada en esta tesis es la siguiente:

Análisis Delictivo es el conjunto de procesos y técnicas de análisis encaminados a proveer información oportuna y pertinente referente a hechos y correlaciones de tendencia delictiva al personal operativo y administrativo durante la planeación de acciones destinadas a prevenir, evitar actividades delictivas, y esclarecer los casos.

#### **2.3.3.3.1. Tipos de Análisis Delictivo**

Existen diversas formas de agrupar los diferentes tipos de análisis que se desarrollan en los departamentos de policía (Martínez 2009 citando a Osborne y Wernicke 2006).

- ❖ Análisis Delictivo Táctico: Basado en determinar dónde, cuándo y cómo se están cometiendo los delitos para informar a los oficiales de policía e investigadores. Sin embargo, otro rol de este tipo de análisis es el encontrar patrones comunes a cierto tipo de delito, que prediga nuevos hechos delictivos dentro de la misma área y modo de actuar de los delincuentes.
- ❖ Análisis Delictivo Estratégico: Desarrollado por los mandos superiores encargados de la de seguridad pública para resolver problemas delictivos de gran impacto y planear proyectos a largo plazo.
- ❖ Análisis Delictivo Administrativo: Centrado en proveer a los diferentes organismos policiacos datos, estadísticas e información referente a la tendencia delictiva presentada dentro de una zona.
- ❖ Análisis Delictivo de Investigación: Se encarga del estudio de la conducta criminal para la creación de perfiles (socio-económicos, culturales y psicológicos), de sospechosos, delincuentes y víctimas, a efecto de detectar posibles delincuentes, asesinos en serie y fundamentar los estudios de victimología.
- ❖ Análisis Delictivo de Inteligencia: Se desarrolla en los niveles federales y estatales, centrándose principalmente en el estudio del crimen organizado, terrorismo, etc.
- ❖ Análisis Delictivo de Operaciones: Tiene como objetivo evaluar el aprovechamiento de los recursos dentro de una organización policiaca (presupuestos, recursos humanos, inventarios de armas, estados de fuerza, aprovechamiento tecnológico, etc.).

#### **2.3.3.4. Clasificación de los Delitos en Perú**

En el país según la Región Policial Lambayeque, los delitos se dividen en 12 clasificaciones los cuales a su vez presentan ciertas variedades y subtipos, los cuales se detallan a continuación:

**Tabla 02: Delito CVCS**

<b>Contra la Vida el cuerpo y la Salud</b>	
<b>Lesiones</b>	Arma blanca Agresión Objeto contundente Ingesta de sustancias tóxicas Quemadura Electrocución Exposición Precipitación Autolesiones por Ingesta de sustancias tóxicas Negligencia Médica Otros
<b>Aborto</b>	Consentido No consentido Autoaborto Con muerte súbita

Fuente: Región Policial Lambayeque 2011.

**Tabla 03: Delito contra la Familia**

<b>Contra la Familia</b>
Atentados contra la patria potestad Omisión asistencia familiar Matrimonio ilegal Delito contra el estado civil Abandono de Hogar Retiro forzado de Hogar

Fuente: Región Policial Lambayeque 2011.

**Tabla 04: Delito contra la Libertad**

<b>Contra la Libertad</b>	
<b>Violación libertad personal</b>	Secuestro típico Secuestro al paso Coacción Trata de personas
<b>Violación de la intimidad</b>	Apoderamiento de documentos personales ajenos. Intercepción de transmisión de datos de carácter personal
<b>Violación de domicilio</b>	Transgresión de domicilio
<b>Violación de Libertad sexual</b>	Violación sexual a menor de edad Violación sexual a mayor de edad Tentativa de violación Acoso Seducción

Fuente: Región Policial Lambayeque 2011.

**Tabla 05: Delito contra el Patrimonio**

<b>Contra el Patrimonio</b>	
<b>Hurto simple y hurto agravado</b>	Domicilios Locales comerciales Centro educativos Otros
<b>Robo simple y robo agravado</b>	Asalto y robo a personas Asalto y robo a entidades
<b>Apropiación Ilícita</b>	Apropiación de bienes Apropiación de propiedad ajena
<b>Estafas y otras defraudaciones</b>	Tarjeta de crédito Seudo profesional Cuento casa propia Cuento VISA Cuento TINKA Cuento cambiazo Cuento terreno, otros.
<b>Daños simples y agravados</b>	Daños materiales en domicilio
<b>Otros</b>	Receptación Usurpación Extorsión

Fuente: Región Policial Lambayeque 2011.

**Tabla 06: Delito contra la Humanidad**

<b>Contra la Humanidad</b>
Desaparición forzada Genocidio Tortura

Fuente: Región Policial Lambayeque 2011.

#### **2.3.3.5. Legislaciones y Políticas en contra la Criminalidad**

##### **2.3.3.5.1. La Constitución y el Sistema Nacional de Seguridad Ciudadana**

La Constitución Política del Perú, en sus dos primeros artículos establece los derechos fundamentales de la persona, las bases de un sistema de seguridad ciudadana son estos derechos, especialmente los relacionados con la vida, la tranquilidad, la libertad y seguridad personales (Verástegui 2005).

Los artículos analizados a continuación definen específicamente las labores que debe ejecutar la policía, cuál es su misión y función, que resulta siendo clave para las políticas de seguridad ciudadana.

**Tabla 07: Artículos de la Constitución para garantizar la Seguridad Ciudadana**

Artículo 166	Artículo 197
<ul style="list-style-type: none"> <li>• La Policía Nacional tiene por finalidad fundamental garantizar, mantener y restablecer el orden interno.</li> <li>• Presta protección y ayuda a las personas y a la comunidad.</li> <li>• Garantiza el cumplimiento de las leyes y la seguridad del patrimonio público y del privado.</li> <li>• Prevenir, investigar y combatir la delincuencia.</li> <li>• Vigilar y controlar las fronteras</li> </ul>	<ul style="list-style-type: none"> <li>• Las municipalidades promueven, apoyan y reglamentan la participación vecinal en el desarrollo local.</li> <li>• Brindar servicios de seguridad ciudadana, con la cooperación de la Policía Nacional del Perú, conforme a ley.</li> </ul>

Fuente: Constitución Política del Perú (1993)

- ❖ **Decretos y Resoluciones Supremas:** En primer lugar, el DS. 105-2002-PCM otorga al problema de seguridad ciudadana rango de Política de Estado dentro del Acuerdo Nacional, eso lo sitúa en primer orden de prioridad nacional para la planificación de políticas públicas. En segundo lugar, la RS. 0965 y el DS. 012-2003-IN están destinados a brindar una adecuada fuerza de seguridad y una estructura de organización de la seguridad que le permita a la sociedad interactuar con la fuerza policial.

**Tabla 08: Decretos para garantizar la Seguridad Ciudadana**

DS. 105-2002-PCM, que institucionaliza al Foro del Acuerdo Nacional	RS. 0965, que crea la Comisión Especial de Reestructuración de la Policía Nacional	DS.012-2003-IN, que aprueba el Reglamento de la Ley del Sistema Nacional de Seguridad Ciudadana.
<ul style="list-style-type: none"> <li>• Para garantizar el objetivo de Democracia y Estado de Derecho, uno de los compromisos del Acuerdo Nacional es Preservar el orden público y la seguridad ciudadana, garantizando que la expresión de nuestras diferencias no afecte la tranquilidad, justicia, libertad de las personas y respeto a la propiedad pública y privada.</li> <li>• Por otra parte se establece: Erradicación de la violencia y fortalecimiento del civismo y de la seguridad ciudadana, para ello el Estado implementará una serie de acciones.</li> </ul>	<ul style="list-style-type: none"> <li>• El Gobierno, con RS.0965-2001-IN de octubre de 2001 creó la Comisión Especial de Reestructuración de la Policía Nacional, con la finalidad de mejorar sustantivamente los servicios policiales, acercar más la Policía a la comunidad y dignificar la función policial mejorando los niveles de vida del personal.</li> </ul>	<ul style="list-style-type: none"> <li>• Tanto la Ley como el Reglamento del Sistema Nacional de Seguridad Ciudadana, establecen 4 instancias integrantes del sistema:               <ul style="list-style-type: none"> <li>a) Consejo Nacional de Seguridad Ciudadana (CONASEC)</li> <li>b) Comités Regionales de Seguridad</li> <li>c) Los comités Provinciales de Seguridad Ciudadana</li> <li>d) Los Comités Distritales de Seguridad Ciudadana.</li> </ul> </li> </ul>

Fuente: Verástegui (2005), cita decretos establecidos por el Presidente Constitucional de la República 2002.



#### **2.3.3.5.2. Nuevo Código Procesal Penal - NCPP**

El nuevo modelo de administración de justicia penal en el Perú, vigente desde el 01 de Abril del año 2006, cambió radicalmente el sistema pasándolo de Inquisitivo a Acusatorio, por ende los métodos y procedimientos utilizados para llevar a cabo la Investigación Criminal, también cambiaron, siendo una de las principales reglas o principios la afirmación de la libertad, y no como antes cuando la detención era la regla y la libertad una excepción (Müller 2006).

En el modelo Acusatorio, se produce un cambio radical de la metodología de la Investigación Criminal y al desaparecer la investigación previa realizada por la Policía Nacional del Perú, conjuntamente con el Atestado Policial, desaparece también el método policial de investigación criminal utilizado tradicionalmente; el cual corresponde ahora a los Fiscales, frente a la comisión de un delito, obtener los elementos de convicción necesarios para la acreditación de los hechos delictivos, así como para identificar a los autores o partícipes en su comisión (Art. 65º, inciso 1 del NCPP).

El NCPP señala que el Fiscal en cuanto tenga noticia del delito, debe realizar las primeras Diligencias Preliminares o dispondrá que las realice la Policía Nacional (Art. 65º, inciso 2 del NCPP).

Además el NCPP indica que la función de investigación de la Policía está sujeta a la conducción del Fiscal y que debe tomar conocimiento de los delitos y dar cuenta inmediata al Fiscal, sin perjuicio de realizar las diligencias de urgencia e imprescindibles para impedir sus consecuencias, individualizar a sus autores y partícipes, reunir y asegurar los elementos de prueba que puedan servir para la aplicación de la ley penal (Art. 67º del NCPP)

En el nuevo modelo acusatorio, el Fiscal investiga el delito, reúne las pruebas necesarias y decide si formula o no la acusación escrita; garantizándose los derechos del procesado, como el que se presuma inocente y no se le restrinja la libertad, salvo que ella constituya peligro procesal de que se sustraiga a la acción de la justicia o se perturbe la actividad probatoria.

En este nuevo contexto, la Criminalística Policial, bajo la conducción del Fiscal tiene una relevante participación contribuyendo al logro de los objetivos de la investigación criminal. Teniendo en cuenta 3 áreas importantes, la primera es la participación de los peritos policiales en el estudio de la Escena de Delito, la cual tiene como objetivo verificar el hecho, recolectar datos y testimonios útiles así como tener una apreciación

preliminar reconstructiva del caso; la segunda es el trabajo en el Laboratorio de Criminalística, donde el proceso criminalístico pericial convertirá los indicios y evidencias acopiados en la escena o aquellos recepcionados en la fase preliminar investigatoria, en informes o dictámenes periciales; y la tercera es la Identificación, para demostrar que una persona es aquella que se busca (Müller 2006).

Mediante la actividad investigativa, los policías construyen las pruebas que el Ministerio Público ofrecerá ante los jueces y que serán objeto de controversia por los defensores durante el juicio. Ello exige una racionalidad policial moderna, además de cumplir la legalidad en sus procedimientos y prácticas, interiorice el nuevo modelo procesal penal bajo el sistema de libertad de prueba, donde ésta última es entendida como el conjunto de argumentos que pueden sustentarse con elementos introducidos al proceso durante la etapa de investigación que sirven para descubrir la verdad de los hechos alegados en la acusación (Díaz 2007)

La investigación criminal ha cambiado de paradigma: Las actividades investigativas que realizan la Fiscalía y la Policía e igualmente la consecución de los elementos materiales, no contempla pruebas por si solas mientras no sean presentadas y debatidas en el juicio público, en que mediando el derecho a la defensa y a la contradicción, sean sometidas a la apreciación crítica del juez, quien no interviene en la investigación y debe ser verdaderamente imparcial frente a la presentación que las partes hagan sobre el caso (Domínguez s.f.).

La puesta en marcha de un nuevo sistema de justicia penal en el país impone un cambio de paradigma en la labor policial, particularmente en el ámbito de la Investigación Criminal. Este nuevo paradigma, busca que la evidencia se demuestre científicamente, a través de pruebas obtenidas de manera técnica y científica, es decir la prueba en el nuevo modelo penal no consiste en averiguar, sino en verificar.

#### **2.3.3.6. La Investigación Criminal aplicada actualmente en la Policía Nacional del Perú**

Müller (2006), sostiene que el impacto que ha tenido el nuevo código procesal penal ha sido grande, por un lado los pesquisas policiales, simplemente se han quedado sin trabajo, por otro lado los Laboratorios de Criminalística decaen ante la falta de personal especializado y de recursos e insumos necesarios para poder contribuir a los objetivos del nuevo modelo procesal penal.

La criminalidad avanza incontenible y no faltan experimentados investigadores que culpan de este incremento a la aplicación del nuevo sistema penal. La capacitación ha sido insuficiente y en muchos casos ha sido nula, la mayoría de Policías solo entienden que se viene produciendo una interferencia en sus funciones por parte del Ministerio Público.

En la práctica se observa que existe una total resistencia de orden mental en los Policías que cumplen funciones de investigación, al no poder documentar el resultado de la investigación criminal que se les encomienda a través de los Fiscales, en un documento similar al Atestado Policial, conforme estaban habituados.

El documento que se elabora de acuerdo a las nuevas prácticas es el Informe Policial, el cual no es un documento del resultado de las investigaciones, sino de las diligencias practicadas con un ligero análisis de las mismas que le sirven al Fiscal para decidir si apertura o no la etapa de la Investigación Preparatoria.

En resumen, en este nuevo contexto, la Policía Nacional debe realizar una tarea investigativa Técnico-Científico. Recoger en el lugar del hecho la obtención de todas las evidencias, indicios, huellas, rastros, etc. que con un tratamiento de rigor técnico científico, posibiliten la reproducción conceptual de los acontecimientos motivo de la investigación.

Realmente es una tarea muy difícil dejar atrás las influencias del modelo inquisitivo y lograr que finalmente la Policía Nacional se integre eficientemente en este gran cambio que impone la aplicación del nuevo modelo procesal penal acusatorio en el PERÚ.

Pero tampoco se puede abandonar a la Institución Policial y que sus integrantes sigan adecuándose a su criterio al nuevo sistema penal, sin dotarla de los recursos suficientes, ni de la capacitación necesaria para internalizar el nuevo paradigma, por tanto se requiere concientizar brindando todas las medidas de apoyo para lograr una renovada Policía Judicial y una flamante Policía Científica.

#### **2.3.4. Metodología**

##### **2.3.4.1. Open UP**

En lo correspondiente al desarrollo del Sistema, se decidió optar por las Metodología Ágiles, para el desarrollo de la herramienta por estar especialmente orientadas a proyectos pequeños, constituyen una solución a medida para ese entorno, aportando una elevada

simplificación que a pesar de ello no renuncia a las prácticas esenciales para asegurar la calidad del producto.

De las cuales se seleccionó OPEN UP, estructura el ciclo de vida de un proyecto en cuatro fases: concepción, elaboración, construcción y transición. El ciclo de vida del proyecto provee a los interesados un mecanismo de supervisión y dirección para controlar los fundamentos del proyecto, su ámbito, la exposición a los riesgos, el aumento de valor y otros aspectos.

Garcilazo (2008) menciona que OpenUP es un marco del proceso del desarrollo del software Open Source que en un cierto plazo, se espera que cubra un amplio sistema de necesidades para los proyectos de desarrollo.

**Fig. 04: Ciclo de Vida del Proyecto**



#### **FASE 01: CONCEPCIÓN**

El propósito en esta fase es lograr concurrencia entre todos los interesados sobre los objetivos del ciclo de vida para el proyecto.

Hay cuatro objetivos para la fase de Inicio que clarifican el alcance, los objetivos del proyecto y la viabilidad de la solución proyectada:

- *Entender qué construir.* Determine la Visión, el alcance del sistema y sus límites. Identifique quién está interesado en este sistema y por qué. Es decir en esta etapa se delimitará el alcance del Sistema de Alerta Temprana, se determinan todas las características que va a presentar y se identifican a los interesados (comisarias, Región Policial Lambayeque).
- *Identifique la funcionalidad clave del sistema.* Decida qué requerimientos son los más críticos; se determina la funcionalidad del Sistema de Alerta Temprana.
- *Determine al menos una posible solución.* Se identificará al menos una arquitectura candidata y su viabilidad para el Sistema de Alerta Temprana.
- Entienda el costo, el cronograma y los riesgos asociados al desarrollo del proyecto.

## **FASE 02: ELABORACIÓN**

El propósito general de esta fase es mitigar riesgos técnicos y no – técnicos.

Hay objetivos para la fase de Elaboración que le ayudan a direccionar los riesgos asociados con los requisitos, la arquitectura, los costos y el cronograma:

- Obtenga un entendimiento más detallado de los requisitos. Se tendrá un buen entendimiento de la mayoría de requisitos que le permitan crear un plan más detallado y obtener ganancia de los interesados. Asegúrese de ganar profundidad en el entendimiento de los requisitos más críticos a ser validados por la arquitectura.
- Diseñar, implementar, validar y establecer la línea base para la arquitectura. Diseñe, implemente y pruebe un esqueleto estructural del Sistema de Alerta Temprana. Aunque la funcionalidad no sea completa aún, muchas de las interfaces entre los bloques de construcción son implementadas y probadas. Esto se refiere a una arquitectura ejecutable.
- Mitigar los riesgos esenciales y producir un cronograma exacto y unos costos estimados. Muchos riesgos técnicos son dirigidos como un resultado de detallar los requisitos y de diseñar, implementar y probar la arquitectura. Refine y detalle el plan de proyecto de alto nivel.

## **FASE 03: CONSTRUCCIÓN**

El propósito de esta fase es completar el desarrollo del Sistema de Alerta Temprana basado en la arquitectura.

Hay objetivos para la fase de Construcción que nos ayudan a tener un desarrollo con costo-eficiente de un producto completo - una versión operativa del sistema - que pueda ser entregada a la comunidad de usuarios:

- Desarrolle iterativamente un producto completo que esté listo para hacer transición a su comunidad de usuarios. Describa los requisitos restantes, complete en detalles los diseños, complete la implementación y prueba del software. Libere la primera versión operativa del software (beta) del Sistema de Alerta Temprana y determine si los usuarios están listos para que la aplicación sea desplegada.
- Minimice el costo de desarrollo y alcance algún grado de paralelismo. Optimice los recursos y promueva el paralelismo de desarrollo entre desarrolladores o equipos de desarrolladores,

por ejemplo, asignar componentes que puedan ser desarrollados independientemente una del otro.

#### **FASE 04: TRANSICIÓN**

El propósito de esta fase es asegurar que el Sistema de Alerta Temprana está listo para ser distribuido a los usuarios.

Hay objetivos para la fase de Transición que le ayudan a afinar elegantemente la funcionalidad, el desempeño y la calidad total de la versión beta del Sistema de Alerta Temprana desde el final de la fase previa:

- La prueba beta valida que las expectativas del usuario sean satisfechas. Esto típicamente requiere algunas actividades de afinamiento, tales como depuración de errores y mejora del desempeño y la usabilidad.
- Lograr la concurrencia de interesados que el despliegue se ha completado. Esto puede implicar varios niveles de pruebas para la aceptación del producto, incluyendo pruebas formales e informales y las pruebas beta.
- Mejorar el desempeño en futuros proyectos a través de lecciones aprendidas. Documentar las lecciones aprendidas y mejorar el ambiente de los procesos y las herramientas para el proyecto.

Se utilizó el Lenguaje de Programación Visual Basic .NET, la herramienta de desarrollo Microsoft Visual Studio 2008 y el Sistema Gestor de Base de Datos Microsoft SQL Server 2005.

#### **2.3.4.2. Crisp-DM**

Para el desarrollo de la investigación se tomó en cuenta la metodología CRISP-DM (Cross- Industry Standard Process for Data Mining), es una metodología de libre distribución utilizada para proyectos de Minería de Datos, desarrollada en 1999 por el consorcio de empresas europeas, Pete Chapman y Randy Kerber (NCR - Dinamarca), AG (Alemania), Julian Clinton, Thomas Khabaza y Colin Shearer (SPSS - Inglaterra), OHRA (Holanda) y, Thomas Reinartz y Rüdiger Wirth (DaimlerChrysler)

Esta metodología consta de seis fases (Hernández, Ramírez y Ferri 2004):

#### **FASE 01: COMPRENSIÓN DEL NEGOCIO**

Objetivos y requerimientos desde una perspectiva no técnica.

- ❖ Establecimiento de los objetivos del negocio (Objetivos, necesidades y criterios de éxito)
- ❖ Evaluación de la situación (Requerimientos, supuestos, restricciones)
- ❖ Generación del plan del proyecto (plan, herramientas, equipo y técnicas)

## **FASE 02: COMPRENSIÓN DE LOS DATOS**

Familiarizarse con los datos teniendo presente los objetivos del negocio

- ❖ Recopilación inicial de datos
- ❖ Descripción de los datos
- ❖ Exploración de los datos
- ❖ Verificación de calidad de datos

## **FASE 03: PREPARACIÓN DE LOS DATOS:**

Obtener la vista minable o dataset

- ❖ Selección de los datos
- ❖ Limpieza de datos
- ❖ Construcción de datos
- ❖ Integración de datos
- ❖ Formateo de datos

## **FASE 04: MODELADO**

Aplicar las técnicas de minería de datos a los dataset

- ❖ Selección de la técnica de modelado
- ❖ Diseño de la evaluación
- ❖ Construcción del modelo
- ❖ Evaluación del modelo

## **FASE 05: EVALUACIÓN**

De los modelos de la fase anterior para determinar si son útiles a las necesidades del negocio

- ❖ Evaluación de resultados
- ❖ Revisar el proceso
- ❖ Establecimiento de los siguientes pasos o acciones

## **FASE 06: IMPLEMENTACIÓN**

Explotar utilidad de los modelos, integrándolos en las tareas de toma de decisiones de la organización.

- ❖ Plan de implementación
- ❖ Planificación de la monitorización y del mantenimiento
- ❖ Generación de informe final
- ❖ Revisión del proyecto

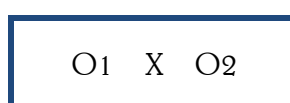
### III. MATERIALES Y MÉTODOS

#### 3.1. Diseño de la Investigación

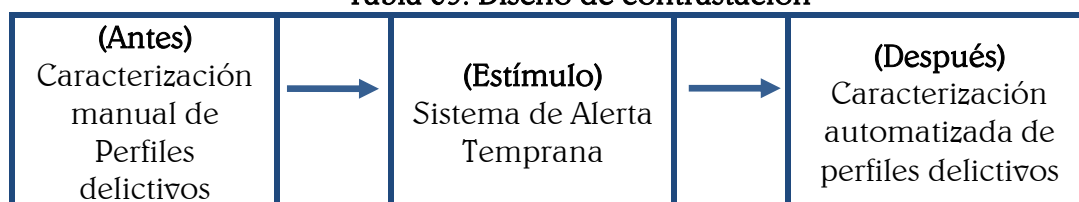
La investigación por ser de carácter aplicada, se utilizó para la contrastación de la hipótesis correspondiente, el diseño en Sucesión (Tabla 09); donde se tuvo la participación de un grupo experimental, el cual fue evaluado y analizado antes y después de haber aplicado el estímulo (Sistema de Alerta Temprana para detectar Patrones Delictivos).

Las conclusiones fueron establecidas por comparación entre la situación antes de aplicar la variable estímulo, y la situación después de aplicar esta variable.

Diseño con pre-test post-test y grupo único.



**Tabla 09. Diseño de contrastación**



#### 3.2. Población

La población objeto de estudio estuvo conformado por los miembros de la Región Policial Lambayeque, los cuales fueron un total de 10 personas y por los registros de las denuncias recogidas por la Región Policial Lambayeque durante el año 2009, se eligió dicha población pues se necesita de las personas para lograr el levantamiento de información y la constante colaboración debido a que ellas utilizaran el sistema de Alerta Temprana y a los registros de las denuncias para recabar información del historial delictivo de cada persona y en base a sus características definir posibles perfiles delictivos.

El total de denuncias en el año 2009 fue de 5376 registros, distribuidas de la siguiente manera.

**Tabla 10: Población de registro de denuncias en 2009.**

Mes	Número de denuncias
Enero	442
Febrero	569



Marzo	523
Abril	618
Mayo	656
Junio	583
Julio	593
Agosto	438
Septiembre	389
Octubre	174
Noviembre	339
Diciembre* (Considera hasta el 10 de diciembre como el cierre del año)	52

### 3.3. Muestra

La muestra seleccionada fue en base al muestreo realizado el cual consideró como elementos de la muestra a un conjunto de 359 registros de la Región Policial Lambayeque, sin embargo para las personas por considerarse un número reducido se tomará en cuenta a las 10 personas que laboran en la Región Policial, es decir esta muestra fue elegida porque existían condiciones previamente contenidas para el acceso al escenario y la interacción con los informantes.

### 3.4. Muestreo

Dada que nuestra población es finita (5376 registros), se utilizó la siguiente fórmula para calcular el tamaño de la muestra.

$$n = \frac{N * Z^2 p * q}{E^2 * (N - 1) + Z^2 * p * q}$$

Donde se considera:

N	Total de la población
n	Tamaño de la muestra
$Z_a^2$	Nivel de fiabilidad de 95% (valor estándar es 1.96 <sup>2</sup> )
P	Probabilidad de que el evento ocurra (en este caso 50% = 0.5)
Q	Probabilidad de que el evento ocurra : 1 – p = 1-0.5 = 0.5)
E	Margen de error de 5% (valor estándar 0.05).

$$n = \frac{5376 * 1.96^2 * 0.5 * 0.5}{0.05^2 * (5376 - 1) + 1.96^2 * 0.5 * 0.5}$$

$$n = \frac{5163.11}{14.40}$$

$$n = 358.60$$

$$n = 359 \text{ registros}$$

El resultado indica un total de 359 denuncias a ser utilizadas, sin embargo para mantener el equilibrio de las denuncias mensuales se consideró óptimo considerar el mismo número de denuncias por mes que es un promedio de 30 denuncias mensuales.

### 3.5. Variables

**3.5.1. Variable Dependiente:** Proceso de detección de patrones delictivos.

**3.5.2. Variable Independiente:** Sistema de Alerta Temprana basado en la minería de datos.

### 3.6. Indicadores

VARIABLE	DIMENSIÓN	INDICADOR	DESCRIPCIÓN
Sistema de Alerta Temprana basado en la minería de datos	Acceso al Sistema de Alerta Temprana	Índice de seguridad de acceso al sistema	Indica el porcentaje de personas de la Región Policial Lambayeque que tienen acceso al Sistema de acuerdo a los niveles de seguridad.
	Características que presenta el Sistema de Alerta Temprana	Índice de satisfacción por el uso del Sistema de Alerta Temprana	Indica el porcentaje de personas de la Región Policial Lambayeque que se sienten satisfechos con el uso del Sistema de Alerta Temprana.
		Usabilidad del Sistema de Alerta Temprana	Indica si el Sistema de Alerta Temprana es fácil de aprender y de utilizar por las personas.
		Tiempo de Respuesta por consulta	Tiempo de demora del sistema para dar una respuesta ante una petición de usuario
		Número de reportes por personas denunciadas	Indica la cantidad de reportes emitidos por el sistema de las personas que han sido denunciadas.
		Tiempo de dedicación en el registro de la denuncia	Tiempo dedicado al proceso de registro de una denuncia.
		Tiempo en búsquedas de información por criterios	Indica el tiempo que se demora el sistema en determinar una persona a un perfil delictivo.
Proceso de detección de patrones delictivos	Denuncia	Tiempo de demora en la emisión resolución de las denuncias	Tiempo de demora en dar respuesta a una denuncia.
		Índice de errores en coincidencias de datos del implicado por expediente	Indica el porcentaje de denuncias registradas con errores (duplicidad e inconsistencia de datos)
	Materiales de escritorio	Costo por materiales de escritorio	Indica la cantidad de material utilizado en el proceso de registro de denuncias.

### 3.7. Técnicas e Instrumentos de recolección de datos

Para la obtención de la información se aplicaron encuestas y entrevistas estructuradas y libres a los miembros de la Región Policial Lambayeque, y a los miembros de la Comisaría del Norte, para conocer la situación actual del proceso que realiza para el control de denuncias y búsquedas de patrones delictivos y desarrollar el Sistema de Alerta Temprana para detectar Patrones Delictivos que permita llevar el registro de las denuncias de manera automatizado y detectar los posibles perfiles delictivos.

#### ❖ **La Técnica del análisis documental**

Se utilizó como instrumentos de recolección de datos las fichas textuales y de resumen; teniendo como fuentes libros, informes y otros documentos, las que permitieron recopilar toda clase de información Teórica– Científica, para estructurar las bases Teóricas y orientar con eficacia la investigación.

Se analizó las actas de registro, informes finales y los diferentes documentos, que nos brindó el personal encargado de llevar el control de los registros de denuncias.

#### ❖ **La Técnica de la Encuesta**

Se utilizó como instrumentos los Cuestionarios, que fueron aplicados a los miembros de la Región Policial Lambayeque, y a los miembros de la Comisaría del Norte de Chiclayo, en febrero del 2011 los cuales hicieron uso de la herramienta, para poder determinar el nivel de Satisfacción que se había logrado alcanzar con la Implementación del Sistema (**Ver Anexo 05**).

#### ❖ **La Técnica de la Entrevista**

Se utilizó como instrumento la guía de entrevista para el SOS PNP Juan Díaz Villegas en cuanto al proceso de Registro de denuncias y definición de Perfiles Delictivos en febrero del 2011 (**Ver Anexo 04**).

#### ❖ **Ficha de Evaluación de la Herramienta**

Se utilizó este instrumento para realizar una evaluación objetiva del Sistema de Alerta Temprana, se realizó la evaluación a los policías involucrados con el registro de denuncias que hicieron uso del sistema en la Región Policial Lambayeque, en mayo del 2012.

Además se evaluó a todos los policías involucrados con el registro de denuncias en la Región Policial Lambayeque y se tomaron los tiempos de demora de registro de denuncia, denominado Pre –Test.

#### • **Pre – Test**

Prueba dirigida a los todos policías (10) involucrados con el registro de denuncias en la Región Policial Lambayeque, sobre el tiempo que demoran en registrar una denuncia; se realizó antes de la implementación de la herramienta.

Después se hizo uso del estímulo, es decir tomaron los tiempo nuevamente pero haciendo uso del Sistema de Alerta Temprana, como una herramienta de apoyo a los policías.

Después de que se interactuó con la herramienta se les aplicó un Post – Test.

- **Post – Test**

Prueba dirigida a los policías (10) involucrados con el registro de denuncias en la Región Policial Lambayeque, sobre el tiempo que demoran en registrar una denuncia; se realizó después de haber hecho uso de la herramienta.

Finalmente se hizo una comparación del antes y después de la implementación del Sistema de Alerta Temprana para identificar patrones delictivos y percibir los cambios en cuanto a la reducción de tiempos en la realización de tareas.

**Tabla 11: Técnicas para la Recopilación de Datos**

<b>Técnica</b>	<b>Instrumento</b>	<b>Observaciones</b>
<b>Entrevista</b>	Guía de entrevista	Constituyó la hoja de ruta de la entrevista.
<b>Encuesta</b>	Cuestionario	Sirvió presentar las preguntas sobre la investigación.
<b>Análisis documental</b>	Guía de análisis Documental	Se utilizó para recopilar la información de textos, libros, revistas, informativos, Internet y otros. Además para recolectar la data (registros de denuncia, cuadros resúmenes, etc.) proporcionada por la Región Policial Lambayeque
<b>Pre – Test y Post – Test</b>	Toma de tiempos	Se usó para medir los tiempos en cuanto al registro y almacenamiento del proceso de registro de denuncia

### 3.8. Técnicas de Procesamiento de datos

Todos los instrumentos que se utilizaron para la recopilación de datos se aplicaron a los 10 policías de la Región Policial de Lambayeque, para lo cual los datos que se obtuvieron fueron procesados y presentados en forma de cuadros, gráficos y resúmenes, mediante la ayuda de la herramienta Microsoft Office Excel 2007, por lo que se aplicó la Estadística Descriptiva, por la que cada cuadro fue analizado exhaustivamente para llegar a las respectivas conclusiones.

Se hizo uso de la Prueba T de Student, aplicada a la toma de tiempos para controlar la demora en los procesos principales antes y después de la implementación del sistema, luego se hizo un análisis comparativo para determinar el nivel de significancia de la implementación propuesta durante la etapa del Post – Test.

### **3.9. Metodología**

En lo correspondiente al desarrollo del Sistema, se decidió optar por la metodología OPEN UP por estar especialmente orientadas a proyectos pequeños, lo cual constituye una solución a medida para ese entorno.

Para el desarrollo de la investigación se tomó en cuenta la metodología CRISP-DM es una metodología de libre distribución utilizada para proyectos de Minería de Datos.

#### IV. RESULTADOS

En esta parte se indican los resultados los cuales corresponden las metodologías tanto de desarrollo como de investigación descritas en el capítulo III. De las cuales la metodología de investigación se encuentra desarrollada dentro de la metodología de desarrollo, la que permitió realizar el trabajo en forma ordenada y consistente pasando por cada una de sus fases.

##### 4.1. Resultado del Pre – Test

El Pre – Test constó en la toma de tiempos para controlar la demora en los procesos principales como del proceso de Registro de denuncia; se obtuvieron los siguientes resultados.

**Tabla 12: Resultados obtenidos en el Pre - Test**

PRE - TEST		
T01	11	minutos
T02	10	minutos
T03	9	minutos
T04	10	minutos
T05	10	minutos
T06	12	minutos
T07	12	minutos
T08	11	minutos
T09	9	minutos
T10	11	minutos

El promedio de registro de una denuncia por parte de los policías es de 10.5 minutos.

##### 4.2. FASE 01: CONCEPCIÓN

Se definieron los criterios principales que permitieron llevar a cabo el desarrollo de la herramienta.

**Entender qué Construir:** Se determinó el alcance del Sistema.

**Tabla 13: Alcance del Proyecto**

ALCANCE
El Sistema de Alerta Temprana, comprende el proceso de registro y almacenamiento de denuncias, determinación de grupos con características semejantes que estuvieron implicados en una denuncia, determinar perfiles delictivos; además permite el mantenimiento de unidad de intervención, delito, tipo de delito, agraviado, implicado y apoderado, también presenta el mantenimiento de usuario (policía), la realización de consultas, la generación de reportes, y finalmente permite evaluar las posibles personas sospechosas ante un hecho que se ha cometido u ocurrido

Luego se identificaron los interesados del Sistema de Alerta Temprana, los cuales fueron:

- Región Policial Lambayeque: Ya que le permitirá incrementar la imagen que tiene en la comunidad chichlayana además de mejorar la calidad de atención y agilizar sus procesos respecto a las denuncias.
- Comisarías: Debido a que va a servirle como una herramienta de apoyo en el proceso de registro y gestión de denuncias.
- Policías: es el principal beneficiario del Sistema, ya que es quien utiliza la herramienta, permitiéndole agilizar sus tareas puesto que el sistema tiene la capacidad de determinar perfiles delictivos, haciendo que el foco de estudio sea más específico y más reducido.

**Identificar la Funcionalidad clave del Sistema:** Se determinaron los requerimientos funcionales del Sistema.

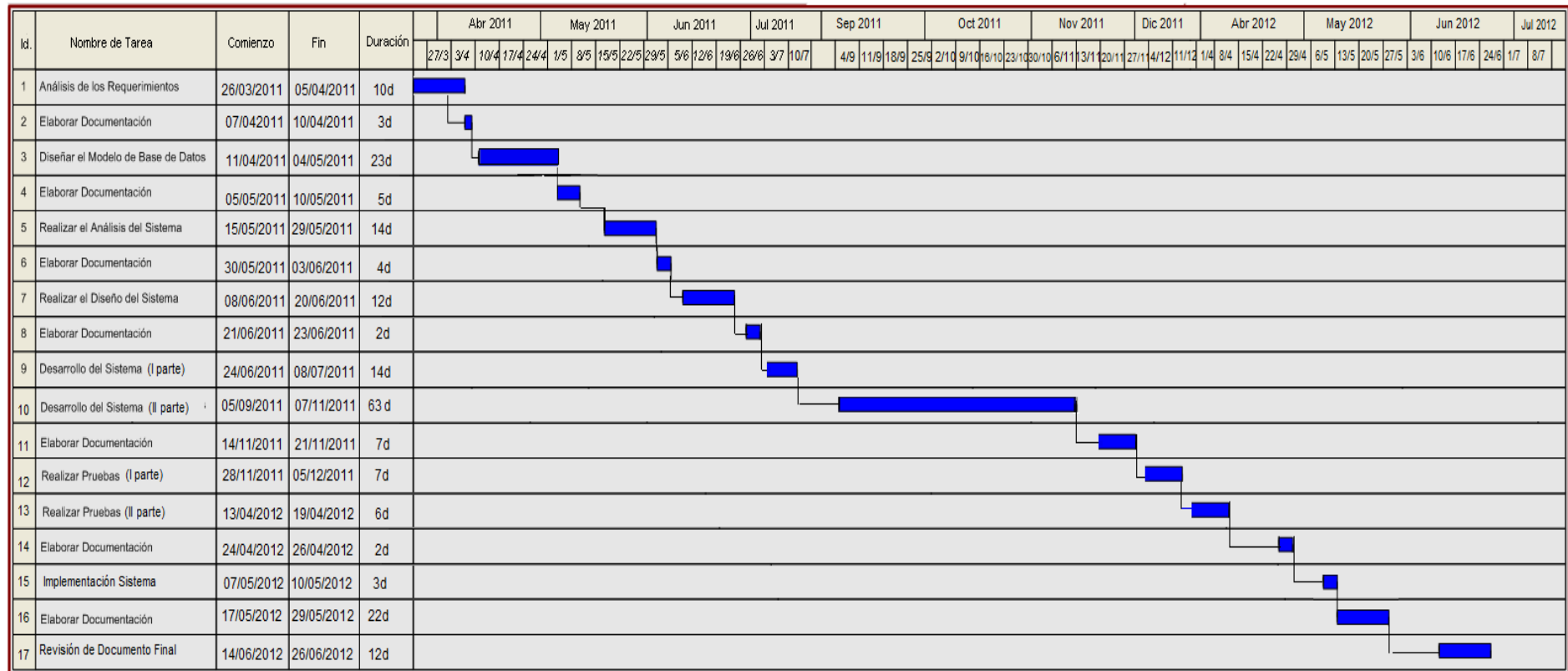
**Tabla 14: Requerimientos Funcionales Clave**

REQUISITOS FUNCIONALES
RF_01: Control de acceso al Sistema por medio de Login
RF_02: Registrar las Denuncias
RF_03: Gestionar los atributos (Unidad de intervención, delito, tipo de delito, agraviado, implicado, apoderado) de las denuncias.
RF_04: Gestionar Usuarios (Policías).
RF_05: Crear Dataset
RF_06: Determinar perfil delictivo (Evaluaciones por cada tipo de perfil delictivo) del implicado
RF_07: Consultar Denuncias registradas por estado o por Comisaría



RF_08: Consultar Denuncias pendientes o en trámite
RF_09: Consultar Denuncias del día
RF_10: Consultar Denuncias del mes
RF_11: Consultar Implicados con más denuncias
RF_12: Consultar Implicados registrados
RF_13: Reporte Implicados registrados
RF_14: Reporte Agraviados registrados

## Determinar Cronograma



### Identificar Costos

Los costos requeridos para el desarrollo fueron financiados por la estudiante, desarrolladora de la tesis Jiménez Berríos Lesly, se detallan a continuación.

**Tabla 15: Costos del Proyecto**

RUBRO	NOMBRE DEL RECURSO		SUBTOTAL
BIENES	Materiales de escritorio		400.00
	Cartuchos de impresión		150.00
	Computadora		1700.00
	Dispositivos de almacenamiento		150.00
	Herramientas Informáticas	Microsoft Visual Studio 2008	280.00
		Microsoft SQL Server 2005 Developer Edition	133.00
		SPSS Statistics 17	108.00
		Matlab R2010a versión para estudiantes	275.00
SERVICIOS	Pasajes		600.00
	Impresiones		250.00
	Anillados		15.00
	Empastado		30.00
	Fotocopiado		25.00
	Luz		753.00
	Internet		1100.00
Total			5969.00

#### 4.2.1. Fase I: Compresión del Negocio

##### 4.2.1.1. Establecimiento de los objetivos del negocio

La Región Policial fue creada con las siguientes finalidades:

- ❖ Realizar estudios e investigaciones estadísticas en el ámbito delictivo.
- ❖ Llevar el control estadístico de las denuncias registradas por todas las comisarías de la ciudad de Chiclayo.
- ❖ Desarrollar y ejercer lineamientos político-criminal llevados.
- ❖ Diseñar programas específicos destinados a actuar sobre la problemática delictiva, tanto preventiva como correctiva, determinando áreas de mayor prioridad, medios idóneos para la producción de las transformaciones que se considere necesarias con relación al tema.

Mediante el cumplimiento de los objetivos del proyecto se logró concluir las siguientes necesidades:

- ❖ Identificar grupos de personas implicadas en un delito, de acuerdo a características comunes.
- ❖ Obtener las reglas que ayudan a comprender los distintos grupos formados.
- ❖ Desarrollar una herramienta que permita caracterizar perfiles delictivos.

Los criterios de éxito del proyecto estuvieron relacionados con el cumplimiento de las necesidades, desde el punto de vista de encontrar patrones de comportamientos en las personas implicadas en un delito a través de la minería de datos.

#### **4.2.1.2. Evaluación de la situación**

La evaluación de la situación actual del problema delictivo se detalló anteriormente en las bases teóricas, específicamente en el capítulo 2.3.3 denominado “Criminalística y Criminología” y en los subcapítulos 2.3.3.4 “Clasificación de los delitos en Perú”, 2.3.3.5 “Legislación y políticas en contra la criminalidad” y 2.3.3.6 “La Investigación Criminal aplicada actualmente en la Policía Nacional del Perú”.

Los requerimientos, supuestos y restricciones se presentan a continuación.

- ❖ Requerimientos: Los diversos requerimientos identificables respecto al proyecto fueron:
  - Definir el alcance del proyecto.
  - Aprender y entender la herramienta de minería de datos a utilizar.
  - Contar con la base de datos del registro de denuncias en las primeras etapas para poder realizar corridas de prueba.
  - Validar los resultados obtenidos en cada una de las pruebas con especialistas en el tema (policías, encargados de la Región Policial).
- ❖ Supuestos: Entre los supuestos nombrados:
  - Al utilizar una base de datos del año 2009, se supuso que el comportamiento de la población implicada en un delito no ha cambiado significativamente desde entonces.
  - Al tratarse de un tema tan complejo, como llegar a entender el comportamiento de una persona que comete un delito, se consideró diversos supuestos que fueron validados por especialistas al momento de analizar los diversos grupos

característicos. Entre ellos se pueden nombrar al supuesto de que muchos de las personas implicadas tienen hijos aun cuando la mayoría es soltera.

Por otro lado, al no tener información sobre la historia de cada persona, se tuvo que suponer que en general esta última no influye de manera fuerte en las acciones de estos individuos.

- ❖ Restricciones: En cuanto a las restricciones, se consideran las siguientes:
  - No se posee información de la persona implicada en un delito acerca de su historial familiar y cultural que ayude a caracterizarlo de una manera más completa en cuanto a patrones de conducta y generación de reglas.
  - Existen en la base de datos campos inconsistentes, con valores erróneos o vacíos que afectan al análisis de los resultados. Esto hace que se tenga que preparar la información previamente llevando forzosamente a la reducción de la muestra poblacional.

#### 4.2.1.3. Generación del plan del proyecto

Se describió el plan para llevar adelante el proyecto, según lo establecido en el cronograma, el cual se detalló con anterioridad. El tiempo total de realización fue de un ciclo académico Marzo – Julio de 2011 y corresponde a una persona.

### 4.3. FASE 02: ELABORACIÓN

**Obtenga un entendimiento más detallado de los requisitos:** Se obtuvo un buen entendimiento de la mayoría de requisitos que permitieron crear un plan más detallado y obtener ganancia de los interesados. Se ganó profundidad en el entendimiento de los requisitos más críticos a ser validados por la arquitectura.

**Tabla 16: Requerimientos Funcionales Detallados**

<b>RF_01: Control de acceso al Sistema por medio de Login</b>	Esta funcionalidad pedirá al usuario (Persona), digitar su nombre de usuario y su contraseña; una vez que se validan y se comprueba que son correctas el usuario podrá acceder a cada una de las opciones del Sistema de Alerta Temprana dependiendo del tipo de usuario (Administrador o usuario común).
<b>RF_02: Registrar las denuncias</b>	Mediante esta característica del Sistema, se pueden registrar todas las denuncias, para que

	puedan ser utilizados en el proceso de definición de perfiles delictivos.
<b>RF_03: Gestionar los atributos (Unidad de intervención, delito, tipo de delito, agraviado, implicado, apoderado) de las denuncias.</b>	Por medio de esta funcionalidad se podrá gestionar (Agregar, Modificar y en algunos casos Eliminar) los atributos presentes en las denuncias.
<b>RF_04: Gestionar Usuarios (Policías).</b>	Permitirá al administrador gestionar (Agregar, Modificar o eliminar) a los usuarios.
<b>RF_05: Crear Dataset</b>	Parte esencial del Sistema de Alerta Temprana, permite definir las características o variables para llevar el proceso de agrupamiento de datos (clustering).
<b>RF_06: Determinar perfil delictivo (Evaluaciones por cada tipo de perfil delictivo) del implicado</b>	Por medio de esta funcionalidad se podrá definir el perfil delictivo de cada persona implicada en una denuncia.
<b>RF_07: Consultar Denuncias registradas por estado o por Comisaría</b>	Permite obtener un listado de todas las denuncias registradas según su estado (Resuelta, En Trámite, Recibida, Archivado y Desistimiento), así mismo permite conocer las denuncias registradas por una Comisaría.
<b>RF_08: Consultar Denuncias pendientes o en trámite</b>	Muestra un listado de todas las denuncias que aún no han sido resueltas.
<b>RF_09: Consultar Denuncias del día</b>	Muestra un listado de todas las denuncias que registradas durante un día específico.
<b>RF_10: Consultar Denuncias del mes</b>	Muestra un listado de todas las denuncias que registradas durante un mes específico.
<b>RF_11: Consultar Implicados con más denuncias</b>	Lista de los implicados o imputados que tienen más denuncias de acuerdo a un rango de fechas.
<b>RF_12: Consultar Implicados registrados</b>	Permite obtener un listado de todas las personas implicadas en una denuncia.
<b>RF_13: Reporte Implicados registrados</b>	Reporte de todas las personas que han sido implicadas en una denuncia
<b>RF_14: Reporte Agraviados registrados</b>	Reporte de todas las personas que han sido agraviadas o han sido víctimas y han registrado una denuncia.

#### 4.3.1. Fase II: Comprensión de los Datos

En esta fase los datos fueron recolectados, organizados, descritos, verificados y limpiados antes de realizar el análisis de los mismos.

##### 4.3.1.1. Recopilación inicial de los datos

La elaboración del proyecto consta de una población conformada por los registros de los delitos cometidos y registrados por la Región Policial Lambayeque durante el año 2009, los cuales son un total de 5376 registros, pero del cual se analizó una muestra de 359 registros.

La información referida a las denuncias fue proporcionada por la Región Policial Lambayeque.

##### 4.3.1.2. Descripción de los datos

La base de datos construida se encuentra desarrollada en SQL Server 2005 y está compuesta por los siguientes datos:

**Tabla 17: Descripción de los Campos de la Base de Datos**

CAMPO	DESCRIPCIÓN	TIPO DE DATO
<b>Tabla: AGRAVIADO</b>		
Idagraviado	Clave principal de identificación del agraviado - víctima	Autonumérico
Nomagraviado	Nombre de la víctima	Texto
Apeagraviado	Apellidos de la Víctima	Texto
Edadagrav	Edad de la Víctima	Numérico
Genero	Sexo de la Víctima	Numérico
Estadocivilagrav	Estado civil de la Víctima	Numérico
Nivelinstruccionagrav	Nivel de Instrucción de la Víctima	Numérico
Ocupacionagrav	Ocupación de la Víctima	Texto
Estado	Estado encontrado de la Víctima	Numérico
Denunciante	En caso de existir más de un agraviado dentro de una denuncia, solo uno representa a las demás personas	Bit

Dniagrav	DNI de la Víctima	Numérico
Idapoderado	Código que hace referencia al apoderado del agraviado	Numérico
<b>Tabla: AGRAVIADO DENUNCIA</b>		
Idagraviado	Clave Principal de identificación del agraviado	Autonumérico
Iddenuncia	Clave Principal de identificación de denuncia	Autonumérico
<b>Tabla: APODERADO</b>		
Idapoderado	Clave Principal de identificación del apoderado	Autonumérico
Nomapoderado	Nombre del apoderado	Texto
Apeapoderado	Apellidos del apoderado	Texto
Edadapoderado	Edad del apoderado	Numérico
Genero	Sexo del apoderado	Numérico
Estadocivil	Estado civil del apoderado	Numérico
Nivelinstruccion	Nivel de Instrucción del apoderado	Numérico
Ocupación	Ocupación del apoderado	Texto
Parentesco	Relación de parentesco con la Víctima	Numérico
Dniapo	DNI del apoderado	Numérico
<b>Tabla: BIEN MATERIAL</b>		
Idbien	Clave Principal de identificación del bien	Autonumérico
Identifpropio	Código propio de identificación del bien	Alfanumérico
Nombrebien	Nombre del bien robado o perdido	Texto
Detallebien	Descripción del bien robado o perdido	Texto
Preciounit	Precio individual por bien robado	Numérico



Comprobante	Para verificar si la víctima cuenta con algún comprobante de propiedad sobre el bien robado	Bit
<b>Tabla: DELITO</b>		
Iddelito	Clave Principal de identificación del delito	Autonumérico
Nomdelito	Nombre del delito	Texto
Idtipodelito	Código que hace referencia al tipo de delito cometido	Numérico
<b>Tabla: DELITO HECHO</b>		
Iddelito	Clave Principal de identificación del delito registrado	Autonumérico
Idhecho	Clave Principal de identificación del hecho que se ha cometido	Autonumérico
Delcircunstancia	Si se infringió en otro delito debido a las circunstancias ocurridas	Bit
<b>Tabla: DENUNCIA</b>		
Iddenuncia	Clave Principal de identificación de denuncia	Autonumérico
estadodenuncia	Estado del proceso de denuncia	Numérico
Destino	Destino de la denuncia	Numérico
Agresión	En caso que la víctima fue lesionada, que tipo de agresión presenta.	Numérico
Fecha	Fecha del registro de la denuncia	Fecha/Hora
Idunidad	El código que hace referencia a la comisaría que registró la denuncia	Numérico
Idhecho	El código que hace referencia al hecho	Numérico

<b>Tabla: HECHO</b>		
Idhecho	Clave Principal de identificación del hecho que se ha cometido	Autonumérico
Nomhecho	Nombre del hecho	Texto
Descripcionhecho	Detalle de lo ocurrido	Texto
Fechorahecho	Fecha y hora de lo ocurrido	Fecha/Hora
Lugarhecho	Lugar o ubicación donde se dio el hecho	Texto
Distrito	Distrito donde ocurrió el hecho	Texto
Tipoarma	Tipo de arma utilizada por el implicado (delincuente)	Numérico
Montorobado	Total monetario sustraído de una víctima.	Numérico
Idobjeto	Código que hace referencia al objeto contundente con que fue atacada o amenazada la víctima	Numérico
Motagresion	Motivo de agresión hacia la víctima	Texto
<b>Tabla: HECHO BIEN</b>		
Idbien	Clave Principal de identificación del bien material robado	Autonumérico
Idhecho	Clave Principal de identificación del hecho que se ha cometido	Autonumérico
Cantidad	Cantidad total de bienes robados	Numérico
<b>Tabla: IMPUTADO</b>		
Idimputado	Clave principal de identificación del implicado	Autonumérico

Nomimputado	Nombre del implicado	Texto
Apeimputado	Apellidos del implicado	Texto
Descripcionfisica	Descripción física del implicado	Texto
Dniimp	DNI del implicado	Numérico
Edadimp	Edad del implicado	Numérico
Sexoimp	Sexo del implicado	Numérico
Estadocivilimp	Estado civil del implicado	Numérico
Condanimicaimp	Condición anímica del implicado en el momento del hecho	Numérico
nivelinstruccionimp	Nivel de Instrucción del implicado	Numérico
Ocupacionimp	Ocupación del implicado	Texto
Reincidenciaimp	Tipo de reincidencia	Numérico
Iddesc	Correlativo de identificador en caso de ser un implicado desconocido NN	Autonumérico
<b>Tabla: IMPUTADO DENUNCIA</b>		
Idimputado	Clave Principal de identificación del implicado	Autonumérico
Iddenuncia	Clave Principal de identificación de la denuncia registrada	Autonumérico
<b>Tabla: INVESTIGA</b>		
Idpoli	Clave Principal de identificación del policía con el cual se va a identificar	Autonumérico
Iddenuncia	Clave Principal de identificación de la denuncia registrada	Autonumérico
<b>Tabla: OBJETO CONTUNDENTE</b>		
Idobjeto	Clave Principal de identificación del objeto contundente con que fue atacada	Numérico

	o amenazada la víctima	
Nomobjeto	Nombre del objeto contundente con que fue atacada o amenazada la víctima	Texto
<b>Tabla: POLI_CATEGORIA</b>		
Idcategoria	Clave Principal de identificación de las categorías existentes respecto al policía	Autonumérico
Nomcategoria	Nombres de las categorías existentes respecto al policía	Texto
<b>Tabla: POLICÍA</b>		
Idpoli	Clave Principal de identificación del policía con el cual se va a identificar	Autonumérico
Nompoli	Nombre del Policía	Texto
Apepoli	Apellidos del Policía	Texto
función	Función que desempeña el Policía	Texto
Usuario	Usuario con el cual podrá acceder al sistema	Texto
Clave	Contraseña con el cual podrá acceder al sistema	Texto
Idcategoria	Código que hace referencia a las categorías que existen respecto al policía	Numérico
Idunidad	Código que hace referencia a unidad de intervención (comisaría)	Numérico
<b>Tabla: TIPODELITO</b>		
Idtipodelito	Clave Principal de identificación del tipo de delito registrado	Autonumérico
Nomtipodelito	Nombre del tipo de	Texto

	delito registrado	
<b>Tabla: UNIDAD INTERVENCIÓN</b>		
Idunidad	Clave Principal de identificación de la comisaría que registra la denuncia	Autonumérico
Nombrequidad	Nombre de la comisaría que registra la denuncia	Texto
Ubicación	Lugar de la comisaría que registra la denuncia	Texto

#### 4.3.1.2.1. Dataset Seleccionado

A priori, los campos o atributos que se eligieron para generar el conjunto de datos (dataset) y efectuar la minería de datos fueron los siguientes:

**Tabla 18: Descripción del Dataset seleccionado**

CAMPO	DESCRIPCIÓN	TIPO DE DATO
<b>Tabla: DELITO</b>		
Iddelito	Clave Principal de identificación del delito	Autonumérico
<b>Tabla: DELITO HECHO</b>		
Delcircunstancia	Si se infringió en otro delito debido a las circunstancias ocurridas	Bit
<b>Tabla: DENUNCIA</b>		
Agresión	En caso que la víctima fue lesionada, que tipo de agresión presenta.	Numérico
<b>Tabla: HECHO</b>		
Fechahecho	Fecha y hora de lo ocurrido	Fecha/Hora
Tipoarma	Tipo de arma utilizada por el implicado (delincuente)	Numérico
Montorobado	Total monetario sustraído de una víctima.	Numérico

<b>Tabla: IMPUTADO</b>		
Idimputado	Clave principal de identificación del implicado	Autonumérico
Nomimputado	Nombre del implicado	Texto
Apeimputado	Apellidos del implicado	Texto
Edadimp	Edad del implicado	Numérico
Sexoimp	Sexo del implicado	Numérico
Estadocivilimp	Estado civil del implicado	Numérico
Condanimicaimp	Condición anímica del implicado en el momento del hecho	Numérico
nivelinstruccionimp	Nivel de Instrucción del implicado	Numérico
Reincidenciaimp	Tipo de reincidencia	Numérico
<b>Tabla: UNIDAD INTERVENCIÓN</b>		
Idunidad	Clave Principal de identificación de la comisaría que registra la denuncia	Autonumérico

El dataset presenta las características de los implicados (datos demográficos), que fueron registradas al momento de realizar una denuncia, así como detalles del hecho que ha cometido y que comisaría ha registrado dicha denuncia, es decir son datos relevantes en cuanto al proceso de registro de una denuncia reducción del ámbito de investigación delictiva.

Los campos que presentan mayores inconsistencias o gran cantidad de valores nulos no se tomaron en cuenta para evitar problemas de calidad de datos, o información no relevante.

#### 4.3.1.2.2. Decodificación de los campos

Para la decodificación de los valores del dataset se tomaron a priori y son los siguientes:

**Tabla 19: Decodificación de la variable Edad del implicado**

Nombre de Variable	Ítem	Descripción
Edadimp	0	N.E
Edadimp	1	0 -18 AÑOS

Edadimp	2	19-20 AÑOS
Edadimp	3	21 -30 AÑOS
Edadimp	4	31 -40 AÑOS
Edadimp	5	41-100 AÑOS

**Tabla 20: Decodificación de la variable Sexo del implicado**

Nombre de Variable	Ítem	Descripción
Sexoimp	0	N.E
Sexoimp	1	Femenino
Sexoimp	2	Masculino

**Tabla 21: Decodificación de la variable Estado civil del implicado**

Nombre de Variable	Ítem	Descripción
Estadocivilimp	0	N.E
Estadocivilimp	1	Soltero
Estadocivilimp	2	Casado
Estadocivilimp	3	Divorciado
Estadocivilimp	4	Conviviente
Estadocivilimp	5	Exconviviente
Estadocivilimp	6	Viudo

**Tabla 22: Decodificación de la variable Condición Anímica del Implicado**

Nombre de Variable	Ítem	Descripción
Condanimicaimp	0	N.E
Condanimicaimp	1	Ebrio
Condanimicaimp	2	Enfermo/alterado
Condanimicaimp	3	Drogado
Condanimicaimp	4	Ecuánime

**Tabla 23: Decodificación de la variable Nivel de Instrucción del implicado**

Nombre de Variable	Ítem	Descripción
nivelinstruccionimp	0	N.E
nivelinstruccionimp	1	Sin instrucción
nivelinstruccionimp	2	Primaria incompleta
nivelinstruccionimp	3	Primaria completa
nivelinstruccionimp	4	Secundaria incompleta
nivelinstruccionimp	5	Secundaria

		completa
nivelinstruccionimp	6	Técnico
nivelinstruccionimp	7	Superior

**Tabla 24: Decodificación de la variable Reincidencia del Implicado**

Nombre de Variable	Ítem	Descripción
reincidenciaimp	0	N.E
reincidenciaimp	1	Primario
reincidenciaimp	2	Reiterante
reincidenciaimp	3	Reincidencia múltiple

**Tabla 25: Decodificación de la variable Incidencia Horaria**

Nombre de Variable	Ítem	Descripción
fechorahecho	1	8:00 A.M-2:00 P.M
fechorahecho	2	2:01 P.M-8:00 P.M
fechorahecho	3	8:01 P.M-2:00 A.M
fechorahecho	4	2:01 A.M - 7:59 A.M

**Tabla 26: Decodificación de la variable Tipo de Arma utilizada durante el desarrollo de un hecho (delito)**

Nombre de Variable	Ítem	Descripción
Tipoarma	0	N.E
Tipoarma	1	Ninguna
Tipoarma	2	Arma blanca
Tipoarma	3	Arma de fuego
Tipoarma	4	Ambas

**Tabla 27: Decodificación de la variable Nombre de Delito**

Nombre de Variable	Ítem	Descripción
Nomdelito	1	Homicidio
Nomdelito	2	Lesiones
Nomdelito	3	Aborto
Nomdelito	4	Atentados contra la patria potestad
Nomdelito	5	Omisión asistencia familiar
Nomdelito	6	Matrimonio Ilegal
Nomdelito	7	Delito contra el estado civil
Nomdelito	8	Violación libertad



		personal
Nomdelito	9	Violación de la intimidad
Nomdelito	10	Violación domicilio
Nomdelito	11	Violación libertad sexual
Nomdelito	12	Proxenetismo
Nomdelito	13	Ofensa pudor público
Nomdelito	14	Hurto simple y hurto agravado
Nomdelito	15	Robo simple y robo agravado
Nomdelito	16	Abigeato
Nomdelito	17	Apropiación Ilícita
Nomdelito	18	Estafas y otras defraudaciones
Nomdelito	19	Fraude en la Administración
Nomdelito	20	Daños simples y agravados
Nomdelito	21	Usurpación / extorción
Nomdelito	22	Delitos monetarios
Nomdelito	23	Acaparam, especul, adulterac.
Nomdelito	24	Contrabando
Nomdelito	25	Falsif. Document. en general
Nomdelito	26	TID
Nomdelito	27	Microcomercialización de drogas
Nomdelito	28	Consumo de Drogas
Nomdelito	29	Tenencia Ilegal de armas
Nomdelito	30	Terrorismo
Nomdelito	31	Desaparición Forzada
Nomdelito	32	Genocidio o Tortura
Nomdelito	33	Violencia Familiar
Nomdelito	34	Retiro Forzado de Hogar

Nomdelito	35	Abandono de hogar
-----------	----	-------------------

**Tabla 28: Decodificación de la variable Unidad de Intervención**

Nombre de Variable	Ítem	Descripción
nombreunidad	1	Cesar Llatas Castro
nombreunidad	2	José Leonardo Ortiz
nombreunidad	3	Campodónico
nombreunidad	4	Del Norte
nombreunidad	5	Atusparia
nombreunidad	6	El Porvenir
nombreunidad	7	De Mujeres y Menores
nombreunidad	8	La Victoria
nombreunidad	9	Quiñones
nombreunidad	10	Región Policial

**Tabla 29: Decodificación de la variable Agresión**

Nombre de Variable	Ítem	Descripción
Agresión	0	N.E
Agresión	1	No hubo
Agresión	2	Física
Agresión	3	Psicológica
Agresión	4	Ambas

**Tabla 30: Decodificación de la variable Delito Circunstancia**

Nombre de Variable	Ítem	Descripción
delcircunstancia	0	N.E
delcircunstancia	1	No hubo
delcircunstancia	2	Lesiones
delcircunstancia	8	Violación libertad personal
delcircunstancia	10	Violación domicilio
delcircunstancia	15	Robo simple y robo agravado
delcircunstancia	20	Daños simples y agravados
delcircunstancia	31	Desaparición Forzada

**Tabla 31: Decodificación de la variable Monto Robado**

Nombre de Variable	Ítem	Descripción
Montrobado	0	N.E
Montrobado	1	No hubo robo
Montrobado	2	Robo valorizado <= 100
Montrobado	3	Robo valorizado > 100 y <= 500
Montrobado	4	Robo valorizado en > 500 y <= 1000
Montrobado	5	Robo valorizado en > 100 y <= 5000
Montrobado	6	Robo valorizado en > 5000

#### 4.3.1.3. Exploración de los datos

Mediante el dataset preliminar seleccionado en el punto anterior, se llevó a cabo la exploración de los datos para familiarizarse con los mismos, y al mismo tiempo poder identificar problemas de calidad.

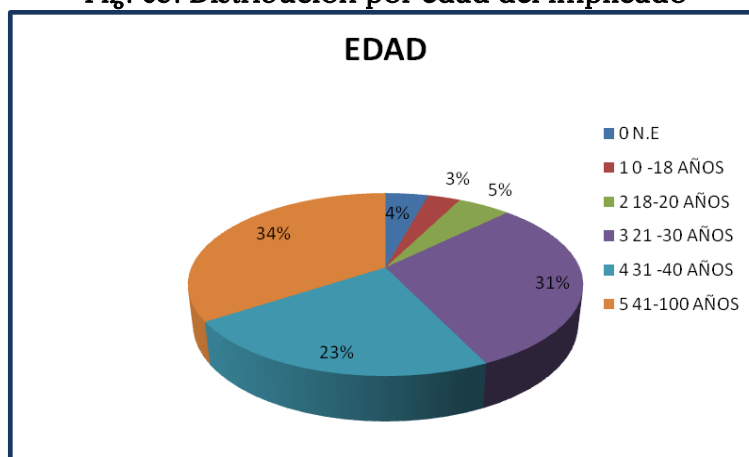
Se comenzó a detectar faltantes de información en algunos campos que se presentan vacíos ("Sin datos"). En esta etapa no se realiza ningún tratamiento de los datos.

[Edad]: Se puede observar que la edad predominante es la entre los 41 y 100 años, seguido por los de 21 a 30 años.

**Tabla 32: Variable Edad del implicado**

EDAD	
DESCRIPCIÓN	FRECUENCIA
0 -18 AÑOS	16
0 -18 AÑOS	26
21 -30 AÑOS	152
31 -40 AÑOS	114
41-100 AÑOS	172
N.E	21
<b>TOTAL</b>	<b>501</b>

Fig. 05: Distribución por edad del implicado

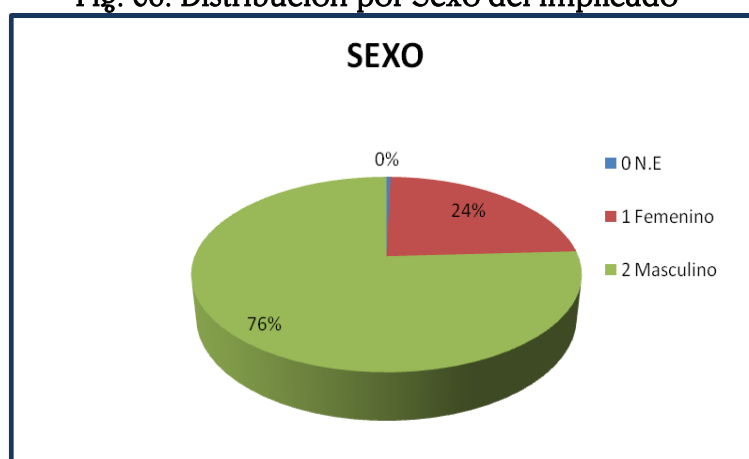


[Sexo]: Se puede observar la distribución del sexo entre los implicados. Se nota claramente que los implicados o denunciados son en su mayoría del sexo masculino.

Tabla 33: Variable Sexo del Implicado

SEXO	
DESCRIPCIÓN	FRECUENCIA
Masculino	380
Femenino	119
N.E	2
<b>TOTAL</b>	<b>501</b>

Fig. 06: Distribución por Sexo del implicado

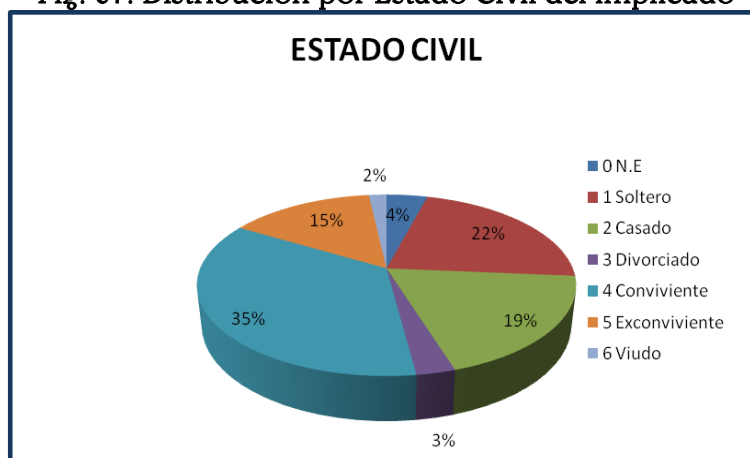


[Estado Civil]: La mayor cantidad de los implicados registrados fueron los convivientes, seguidos de los solteros. Luego, en proporciones similares, le siguen los casados.

**Tabla 34: Variable Estado Civil del Implicado**

ESTADO CIVIL	
DESCRIPCIÓN	FRECUENCIA
Soltero	111
Casado	94
Divorciado	14
Conviviente	178
Ex conviviente	74
Viudo	9
N.E	21
<b>TOTAL</b>	<b>501</b>

**Fig. 07: Distribución por Estado Civil del implicado**

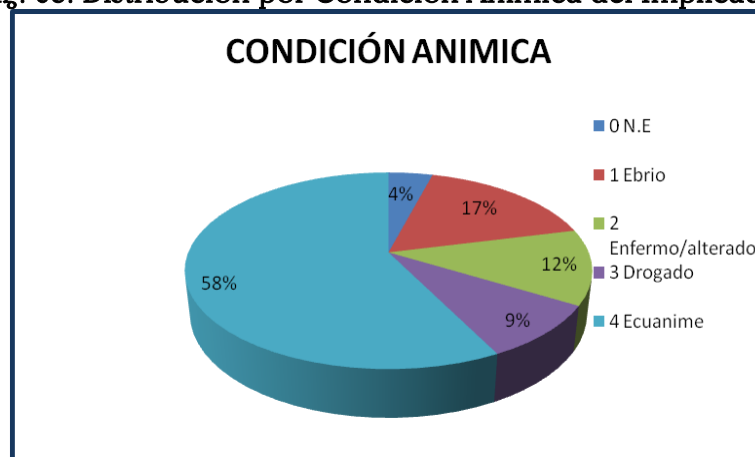


[Condición Anímica]: La mayor cantidad de los implicados al momento de cometer un delito han estado conscientes debido que presentan un estado de ecuanimidad. Luego, en proporciones menores han sido personas que en estado de ebriedad.

**Tabla 35: Variable Condición Anímica del Implicado**

CONDICIÓN ANIMICA	
DESCRIPCIÓN	FRECUENCIA
Ebrio	85
Enfermo/alterado	62
Drogado	45
Ecuánime	288
N.E	21
<b>TOTAL</b>	<b>501</b>

Fig. 08: Distribución por Condición Anímica del Implicado

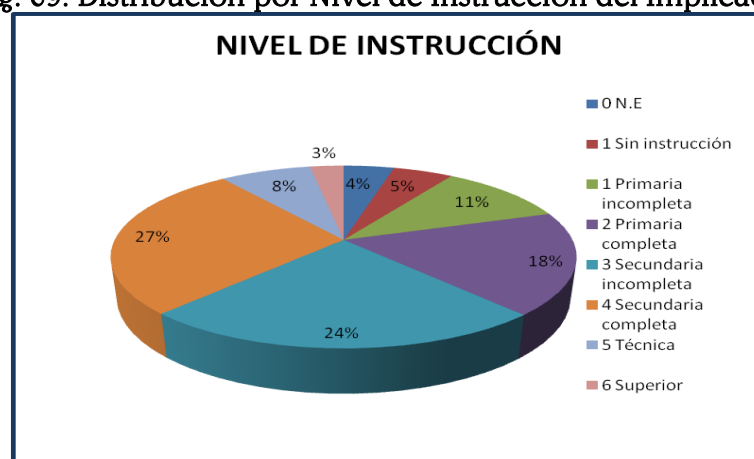


[Nivel de Instrucción]: La mayor cantidad de los implicados solo han logrado llegar al nivel de secundaria completa.

Tabla 36: Variable Nivel de Instrucción del implicado

NIVEL DE INSTRUCCIÓN	
DESCRIPCIÓN	FRECUENCIA
Sin instrucción	25
Primaria incompleta	54
Primaria completa	91
Secundaria incompleta	120
Secundaria completa	138
Técnica	38
Superior	14
N.E	21
<b>TOTAL</b>	<b>501</b>

Fig. 09: Distribución por Nivel de Instrucción del Implicado

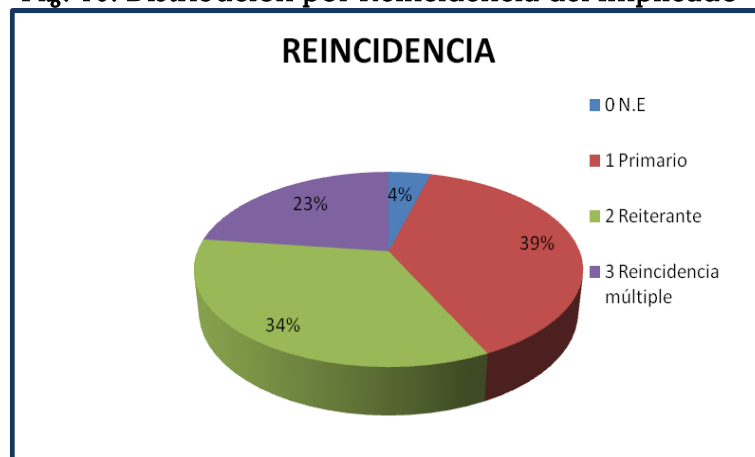


[Reincidencia]: La mayor cantidad de los implicados registrados han tenido una reincidencia primaria, sin embargo existe un gran número de implicados de los cuales en más de una ocasión han cometido el mismo delito

**Tabla 37: Variable Reincidencia del Implicado**

REINCIDENCIA	
DESCRIPCIÓN	FRECUENCIA
Primario	194
Reiterante	171
Reincidencia múltiple	115
N.E	21
<b>TOTAL</b>	<b>501</b>

**Fig. 10: Distribución por Reincidencia del Implicado**

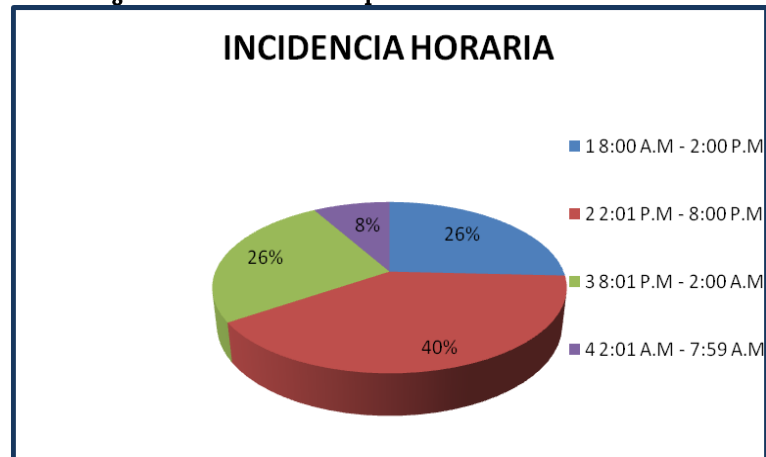


[Incidencia Horaria]: Se puede observar de la distribución de la incidencia horaria que existe una mayor frecuencia de cometer delitos durante la tarde-noche, la cual es seguido con una menor proporción por la mañana.

**Tabla 38: Variable Incidencia Horaria**

INCIDENCIA HORARIA	
DESCRIPCIÓN	FRECUENCIA
8:00 A.M - 2:00 P.M	129
2:01 P.M - 8:00 P.M	200
8:01 P.M - 2:00 A.M	130
2:01 A.M - 7:59 A.M	42
<b>TOTAL</b>	<b>501</b>

Fig. 11: Distribución por Incidencia Horaria

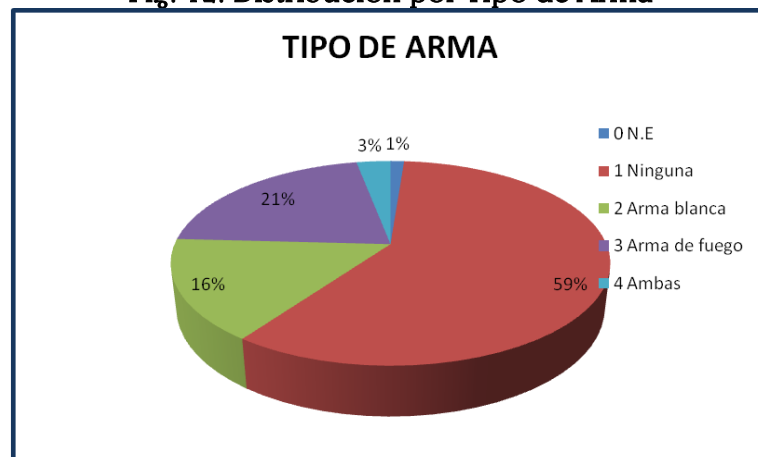


[Tipo de Arma]: Se puede observar de la distribución del tipo de arma, que en la mayoría de delitos o hechos cometidos no se ha utilizado un arma, seguido de esto se encuentra el uso de armas fuego.

Tabla 39: Variable Tipo de Arma

TIPO DE ARMA	
DESCRIPCIÓN	FRECUENCIA
Ninguna	294
Arma blanca	80
Arma de fuego	106
Ambas	15
N.E	6
<b>TOTAL</b>	<b>501</b>

Fig. 12: Distribución por Tipo de Arma



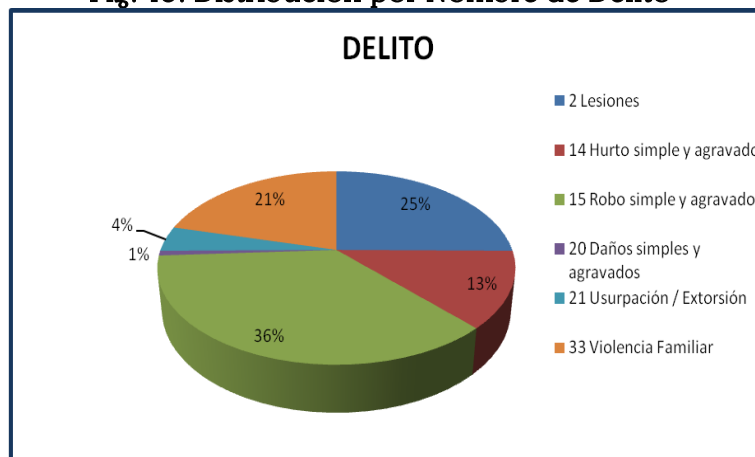


[Delito]: Se puede observar la distribución de los delitos que la mayor cantidad de los delitos registrados se encuentran por Robo simple y agravado seguido de las lesiones.

**Tabla 40: Variable Nombre de Delito**

DELITO	
DESCRIPCIÓN	FRECUENCIA
Lesiones	126
Hurto simple y agravado	64
Robo simple y agravado	181
Daños simples y agravados	4
Usurpación / Extorsión	21
Violencia Familiar	105
<b>TOTAL</b>	<b>501</b>

**Fig. 13: Distribución por Nombre de Delito**



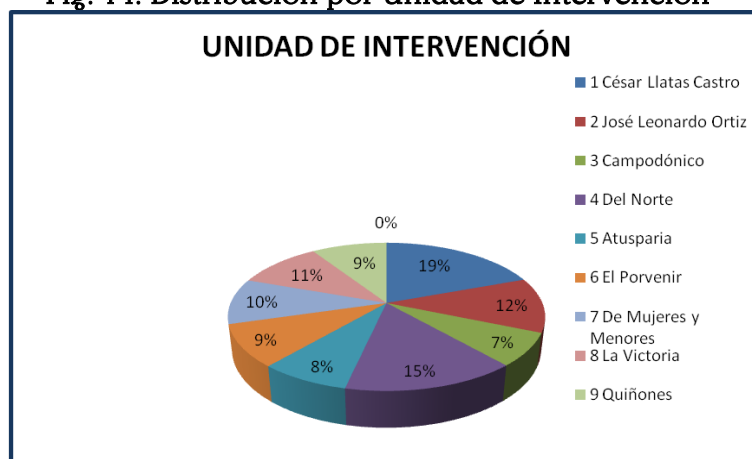
[Unidad de Intervención]: Se puede observar que la comisaría con mayor registro de Denuncias es la principal, es decir la comisaría César Llatas Castro.

**Tabla 41: Variable Unidad de Intervención**

UNIDAD DE INTERVENCIÓN	
DESCRIPCIÓN	FRECUENCIA
César Llatas Castro	97
José Leonardo Ortiz	60
Campodónico	36
Del Norte	75
Atusparia	39

El Porvenir	46
De Mujeres y Menores	50
La Victoria	52
Quiñones	46
Región Policial	0
<b>TOTAL</b>	<b>501</b>

Fig. 14: Distribución por Unidad de Intervención

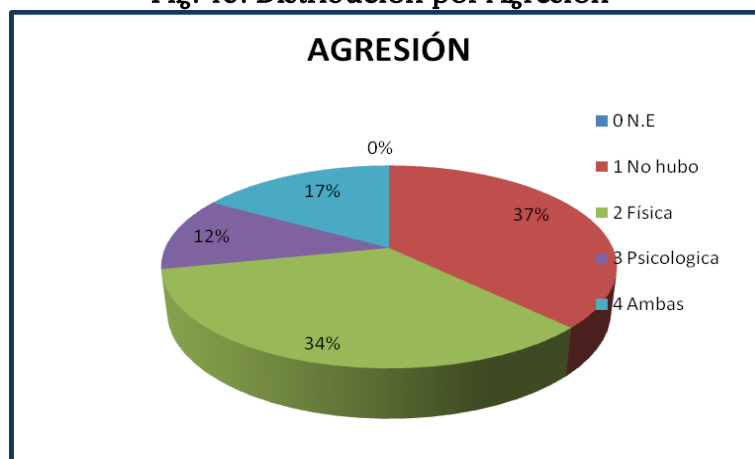


[Agresión]: Se puede observar la distribución de la agresión que en la mayoría de los casos no se agrede a las víctimas en el caso de robo, pero existe una alta cantidad de registros de agresión física que se encuentran en los delitos de maltrato familiar o los que atentan contra la vida el cuerpo y la salud (lesiones).

Tabla 42: Variable Agresión

AGRESIÓN	
DESCRIPCIÓN	FRECUENCIA
No hubo	188
Física	171
Psicológica	59
Ambas	83
N.E	0
<b>TOTAL</b>	<b>501</b>

Fig. 15: Distribución por Agresión

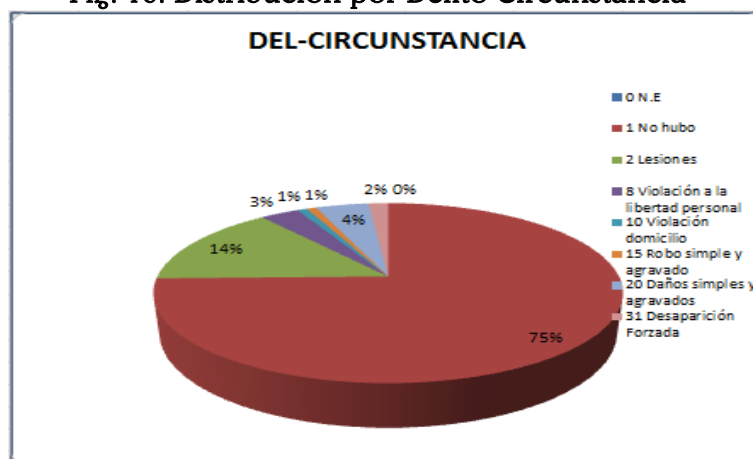


[Delito Circunstancia]: Referido a delitos adicionales cometidos al momento de realizarse un hecho, es decir dado por las circunstancias , como puede ser si se roba ocasionarle lesiones a una persona, en la mayoría de casos registrados no se cometido un delito adicional, pero en proporción menor está el delito de lesiones.

Tabla 43: Variable Delito Circunstancia

DEL-CIRCUNSTANCIA	
DESCRIPCIÓN	FRECUENCIA
No hubo	374
Lesiones	73
Violación a la libertad personal	16
Violación domicilio	4
Robo simple y agravado	4
Daños simples y agravados	22
Desaparición forzada	8
N.E	0
<b>TOTAL</b>	<b>501</b>

Fig. 16: Distribución por Delito Circunstancia

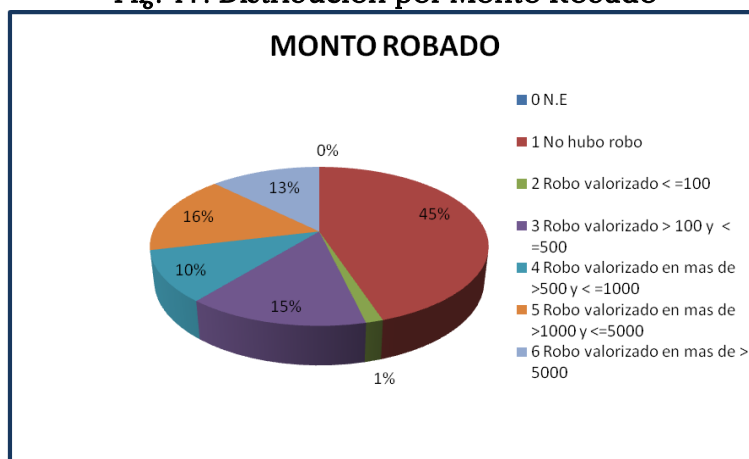


[Monto Robado]: Se puede observar la distribución del monto robado, en caso de haberse cometido un delito, en la mayoría de casos no se ha robado nada, sin embargo cuando si se ha logrado sustraer algún bien la mayoría estuvo valorizado entre 1000 y 5000 soles.

Tabla 44: Variable Monto Robado

MONTO ROBADO	
DESCRIPCIÓN	FRECUENCIA
No hubo robo	225
Robo valorizado < = 100	7
Robo valorizado > 100 y <= 500	73
Robo valorizado en >500 y <= 1000	52
Robo valorizado en >1000 y <= 5000	81
Robo valorizado en > 5000	63
N.E	0
<b>TOTAL</b>	<b>501</b>

**Fig. 17: Distribución por Monto Robado**



#### 4.3.1.4. Verificación de la Calidad de los Datos

Durante la exploración de los datos pudieron observar problemas en la calidad de los mismos, y a su vez, se encontraron varios campos que presentaban datos nulos o ausentes.

En primer lugar, tal como se pudo observar en los gráficos del punto anterior, la totalidad de los mismos presentan registros que tienen valor nulo en alguno de sus campos. La cantidad de veces que aparecen estos campos vacíos varía entre los atributos de acuerdo al momento en que el dato haya sido registrado: Sexo, Edad, Estado civil, Nivel de instrucción, Reincidencia, Condición anímica del implicado. Siendo estos últimos, los que presentan mayores problemas de calidad y ausencia de información.

Comenzando por el campo Sexo, existieron 2 personas no identificadas, es decir se cometió un delito como robo no identificado por el propietario

Con el Edad, Estado civil, existieron 21 implicados registrados que contienen al número cero, es decir no se especificó el promedio de edad de estas personas.

En lo que respecta a la Reincidencia del implicado fueron 21 personas las registradas que no se les especifico el tipo de reincidencia respecto a un delito.

En cuanto al Nivel de instrucción de las personas implicadas también fueron un total de 21 personas registradas sin especificar qué nivel de educación posee, esto se debe por tratarse de personas desconocidas, no identificadas en la realización de un delito (Personas NN).

Tanto la calidad de los datos, como el faltante de los mismos, son cuestiones muy importantes a tener en cuenta previo a realizar la minería en los datos. Esto se analizará en la Fase III Preparación de los Datos.

#### 4.3.2. Fase III: Preparación de los datos

En esta fase se trabajó sobre la preparación del dataset definitivo, la selección y limpieza de los datos. Una vez realizadas estas actividades se procedió a construir los datos a los cuales les fueran necesarios algún tipo de modificación. Es decir, es en esta etapa en donde se preparó todo lo relativo a la base de datos, para poder luego ingresarlas en la herramienta de modelado

##### 4.3.2.1. Vista minable o Dataset

La versión final del dataset que fue utilizado para realizar la minería de datos se presenta en la tabla 45.

**Tabla 45: Dataset Definitivo**

CAMPO	DESCRIPCIÓN	TIPO DE DATO
Edadimp	Edad del implicado	Numérico
Sexoimp	Sexo del implicado	Numérico
Estadocivilimp	Estado civil del implicado	Numérico
Condanimicaimp	Condición anímica del implicado en el momento del hecho	Numérico
nivelinstruccionimp	Nivel de Instrucción del implicado	Numérico
Reincidenciaimp	Tipo de reincidencia.	Numérico
Fechorahecho	Incidencia horaria del delito	Fecha/Hora
Tipoarma	Tipo de arma utilizada por el implicado (delincuente)	Numérico
iddelito	Código del delito	Numérico
idunidad	Código de la comisaría que registró la denuncia	Numérico
Agresión	Tipo de agresión ocasionada	Numérico
Delcircunstancia	Delito según la circunstancia	Numérico
Montrobado	Monto robado	Numérico

#### 4.3.2.2. Selección de los Datos

Los campos que se consideraron para el estudio se eligieron teniendo en cuenta los siguientes factores:

- ❖ Variables que puede tomar cada uno de los campos o atributos.
- ❖ Calidad de los datos.
- ❖ Importancia del atributo para el estudio en cuestión en cuanto a los objetivos del proyecto, información que entrega y significancia del mismo.

Los campos que se consideran para el presente estudio son los siguientes:

***Edad:*** Campo numérico que indica la edad del detenido. Este atributo toma valores desde los 0 años hasta los 100 años.

***Sexo:*** Campo numérico que indica el sexo del detenido.

***Estado Civil:*** Describe el estado civil del implicado en el que se encontraba al momento de cometer un delito.

***Condición anímica:*** Condición anímica del implicado en el momento de cometer un delito.

***Nivel de Instrucción:*** Variable que describe el nivel de educación del implicado

***Reincidencia:*** Representa el tipo de reincidencia del implicado respecto a un delito cometido.

***Incidencia Horaria:*** Esta variable indica el momento en que se cometió el delito, el cual fue especificado por rangos de 6 horas cada grupo.

***Tipo de arma:*** Referido al tipo de arma utilizada por el implicado (delincuente) al momento de cometer un delito.

***Id del delito:*** Variable de identificación del delito que se cometió, los cuales ya han sido establecidos por la PNP.

***Id de la Unidad de Intervención:*** Indica la comisaría que registró la denuncia.

***Agresión:*** Tipo de agresión ocasionada al agraviado al momento de ocurrir un delito.

***Delito Circunstancia:*** Referido a los delitos adicionales cometidos según la circunstancia del desarrollo del hecho principal.

***Monto Robado:*** En el caso de delitos contra el patrimonio u otro tipo de delito en el cual se incurrió en la sustracción de bienes se especificó un rango del valor robado.

#### **4.3.2.3. Limpieza de datos**

Existen registros que son excluidos del análisis por encontrarse alguna inconsistencia, tal como se especificó en el punto **4.2.1.4 “Verificación de Calidad de los Datos”**.

En lo que respecta a las características demográficas (Edad, sexo, reincidencia, nivel de instrucción, etc.) se procedió a eliminar los registros que tenían como valor 0, los cuales no especifican la característica real del implicado.

En segundo lugar, mediante una consulta en el SQL server 2005, se generó la condición de que la base de datos no presente ningún valor nulo para los atributos seleccionados en el dataset. A través de estas modificaciones, se obtuvo una base limpia de inconsistencias y errores.

#### **4.3.2.4. Construcción de datos e Integración de datos**

En la presente etapa se procedió a la construcción de los datos que conformaron el dataset definitivo para su posterior ingreso en la herramienta de modelado.

Se realizó un agrupamiento de los distintos Delitos cometidos a ser evaluados. Como puede observarse en la siguiente tabla (**Tabla 46**), según el tipo de delito, los cuales figuran en la base de datos. Se creyó necesaria esta tarea a fin de facilitar la comprensión de los resultados una vez realizadas las corridas experimentales.



**Tabla 46: Unificación de los Delito**

<b>Contra la Vida el cuerpo y la Salud</b>	Lesiones
	Aborto
<b>Contra la Familia</b>	Matrimonio ilegal
	Delito contra el estado civil
	Atentados contra la patria potestad
	Omisión asistencia familiar
	Abandono de Hogar
	Violencia Familiar
	Retiro forzado de Hogar
<b>Contra el Patrimonio</b>	Hurto simple y hurto agravado
	Robo simple y robo agravado
	Daños simples y agravados
	Usurpación / Extorsión
	Estafas y otras defraudaciones

#### **4.3.3. Fase IV: Modelado**

En esta fase se seleccionó y aplicó las técnicas de modelado.

Previo a seleccionar las técnicas y algoritmos para la explotación de la información, se seleccionó el software a utilizar.

La herramienta informática que se utilizó para llevar a cabo la presente fase fue el SPSS Clementine V17.

##### **4.3.3.1. Selección de la técnica de modelado**

En esta etapa, se seleccionó las técnicas y algoritmos utilizados en el desarrollo de la minería de datos.

Inicialmente se utilizaron técnicas no supervisadas, haciéndose necesario realizar un agrupamiento o clustering de los datos para que el descubrimiento de la información oculta en ellos se torne lo menos difícil posible.

El algoritmo que se utilizó fue K-Means, uno de los métodos más utilizados y el más popular de los métodos de agrupamiento “por partición”, además se tomó en cuenta por ser el algoritmo más utilizado dentro de los antecedentes tomados en la tesis.

El próximo paso fue la validación de los grupos o clúster con los especialistas en temática delictiva. Dicha validación fue de vital importancia para poder seguir adelante con la técnica de clasificación.

Finalmente, se aplicaron Redes neuronales como parte de la técnica de clasificación supervisada teniendo en cuenta los clúster previamente definidos para lograr la identificación de reglas de decisión que ayuden a explicar la composición de cada grupo formado.

##### **4.3.3.2. Construcción del modelo**

Se procedió a explicar cada uno de los modelos utilizados y sus resultados.

###### **4.3.3.2.1. Clustering**

A fin de concretar el agrupamiento de los datos, se utilizó el algoritmo K-Means, obteniéndose 12 grupos de perfiles delictivos a través de sus características más significativas.

La configuración de dicho algoritmo de agrupamiento exige que se elija un número de clúster, para lo cual se utilizó la regla del Codazo, para obtener el número óptimo de clúster, además se eligió el número de iteraciones, para tal caso también se tomó el número que viene por defecto (iteraciones máximas=10) a partir de ello se empezó a realizar una distribución aleatoria inicial y el algoritmo comienza las sucesivas iteraciones, sin embargo el total de iteraciones recorridas fueron 5, momento en la cual el algoritmo dejó de analizar, puesto que los datos ya no mostraban cambios de un clúster a otro.

Asimismo al ser el K-Means un algoritmo que se basa en minimizar la suma del error cuadrático, podría tomarse a ese error como punto de partida para la elección del número óptimo de clúster, para tal caso la regla del codazo, minimizó la probabilidad de error, dando como resultado el número de clúster = 12, así mismo fue el que minimizaba el error cuadrático.

Después de haber realizado la configuración y haber establecido los parámetros, se introdujo el dataset definido anteriormente al modelo de minería de datos para realizar el proceso de clustering correspondiente.

A continuación se muestran los resultados obtenidos.

#### 4.3.3.2.1.1. Resultados del Clustering

El resultado obtenido en el SPSS V17 una vez ejecutado K-Means con 12 clúster, se resume a continuación.

**Tabla 47: Resultados del Proceso de Clustering**

ATRIBUTO	CLÚSTER											
	1	2	3	4	5	6	7	8	9	10	11	12
EDAD	4	3	4	3	4	4	5	4	4	4	3	4
SEXO	2	2	2	2	2	2	2	1	2	2	2	2
ESTADO_CIVIL	3	3	3	3	3	3	4	5	4	3	2	4
NIVEL_INSTRUCCIÓN	6	4	3	4	4	5	3	3	4	4	4	4
REINCIDENCIA	2	1	2	2	2	1	2	2	1	2	2	2
CONDICIÓN_ANIMICA	3	3	4	4	3	3	3	2	2	2	4	4
INCIDENCIA_HORARIA	2	2	2	2	2	2	2	2	3	2	2	3
UNIDAD_INTERVENCIÓN	7	4	6	7	2	4	4	5	7	3	1	2
TIPO_ARMA	2	1	2	2	2	1	3	3	2	1	3	2
AGRESIÓN	4	3	1	1	1	3	4	4	2	3	4	2
DELITO	2	33	15	15	15	33	2	2	2	2	21	21
DELITO_CIRCUNSTANCIA	10	20	20	1	1	1	8	31	1	1	8	1
MONTO_ROBADO	2	3	5	4	4	1	1	1	1	1	6	6
TOTAL	4 (0.8%)	6 (1.3%)	13 (2.7%)	137 (28.5%)	72 (15%)	99 (20.6%)	12 (2.5%)	8 (1.7%)	70 (14.6%)	32 (6.7%)	6 (1.3%)	21 (4.4%)

La tabla anterior resume a los centros de los clúster finales, cuyos valores obtenidos han sido decodificados anteriormente en la *Fase II de la metodología*, denominada *Comprensión de los Datos*.

En primer lugar, y antes de haber analizado los clúster, se encontró en general que la mayoría de los clúster formados muestran que existe una mayor proporción de varones que han cometido un delito (sexo = 2), el cual fue la más frecuente, según los registros obtenidos de la Región Policial y con menor cantidad, es decir con un porcentaje menor, las mujeres:

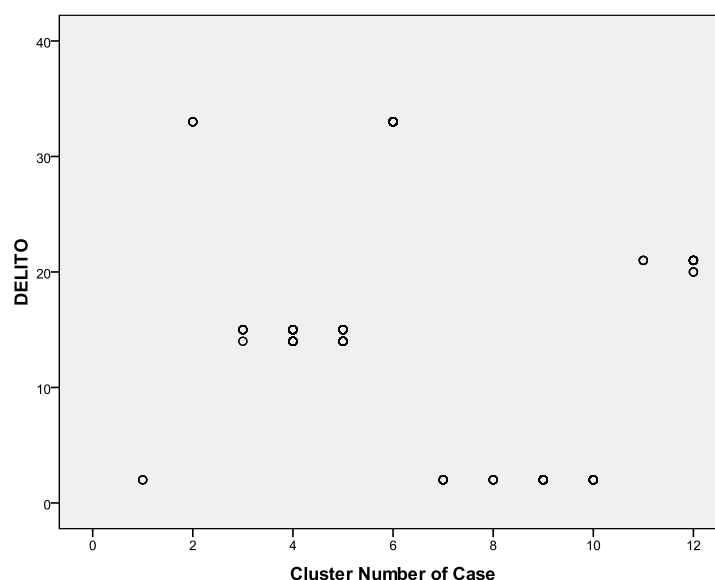
**Tabla 48: Distribución de Reincidencia del Implicado**

REINCIDENCIA	
DESCRIPCIÓN	%
Masculino	75.8
Femenino	23.8
N.E	3.4
TOTAL	100%

Por lo tanto, es de esperar que en todas las corridas de clustering, dentro de los grupos formados se obtenga con mayor frecuencia a los del sexo masculino, puesto que fue el predominante en los registros de denuncias.

### Gráficos de Dispersión para el análisis de clúster

**Graf. 1: Distribución de los clúster según Delito**



El eje de las abscisas estuvo representado por los 12 clústers, mientras que en el de las ordenadas, estuvo representado por los delitos. Dado que el 2 representó al delito de Lesiones, se pudo ver que el cluster 1, 7, 8, 9 y 10 fueron los que conformaron este grupo de delitos, aunque con una mayor proporción por el clúster 9 (55.6%) del total de delitos por lesiones.

Por otro lado el 14 representó al delito de Hurto simple y agravado, y estuvo contenido por los clústers 3, 4 y 5, dentro de estos clústers el de mayor proporción fue el clúster 4 (43.8%) y el de menor proporción fue el clúster 3 (3%) del total de delitos por hurto simple y agravado

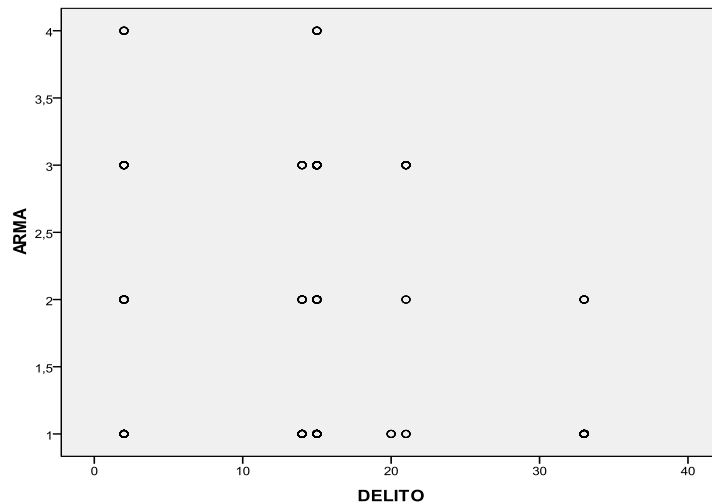
El 15 representó al delito de Robo simple y agravado, y estuvo contenido por los clústers 3, 4 y 5, dentro de estos clústers el de mayor proporción fue el clúster 4 (60.2%) y el de menor proporción fueron los clúster 3 (6.1%) del total de delitos por robo simple y agravado.

El 20 representó al delito de Daños simples y agravados, estuvo contenido por el clúster 12, el cual contuvo dichos delitos (100).

El 21 representó al delito de Usurpación / Extorsión, y estuvo contenido por los clústers 11 y 12, dentro de estos clústers el de mayor proporción fue el clúster 12 (76.2%) y el de menor proporción fue el clúster 11 (23.8%) del total de delitos.

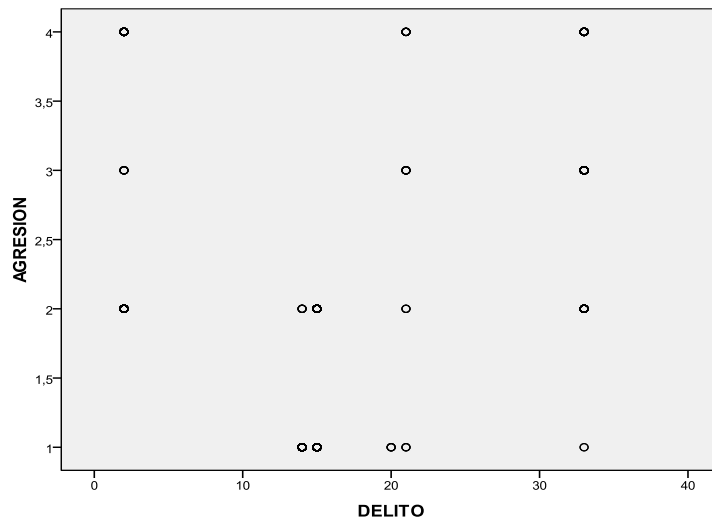
El 33 representó al delito de Violencia familiar, y estuvo contenido por los clústers 2 y 6, dentro de estos clústers el de mayor proporción fue el clúster 6 (94.3%) y el de menor proporción fue el clúster 2 (5.7%) del total de delitos por Violencia familiar.

#### **Graf. 2: Distribución de Tipo de Arma según Delito**



Del gráfico anterior se dedujo que dependiendo el delito se utiliza un tipo de arma, es decir en los delitos 2 (Lesiones), 14 (Hurto simple y agravado) y 15 (Robo simple y agravado) se puede utilizar cualquier tipo de arma sin embargo para el delito 33 (Violencia familiar) no utilizó arma de fuego en el desarrollo del hecho.

**Graf. 3: Distribución de Agresión según Delito**



Del gráfico anterior se dedujo que dependiendo el delito se ocasiona la agresión, es decir en los delitos 2 (Lesiones), 33 (Violencia familiar), 21 (Usurpación / Extorsión) se puede ocasionar cualquier tipo de agresión) física, psicológica o ambas), ya que son delitos que atentan con el bienestar de la persona en los casos de (Hurto o Robo simple y agravado) solo puede ocasionarse una agresión física, puesto que este tipo de delito se cometen rápidamente, y finalmente existen delitos contra el patrimonio que solo ocasionan daños en un inmueble y no hay agresiones hacia las personas como el delito 20 (Daños simples y agravados).

#### 4.3.3.2.1.2. Interpretación de los Clúster

En base a la información que surgió tanto del análisis de los gráficos anteriores como de la investigación y colaboración de los especialistas, se dió una primera interpretación de cada cluster formado.

Fue conveniente destacar, según lo conversado con los especialistas, que para poder realizar un análisis más profundo y exacto, habría que tener acceso a la historia individual de cada uno de los implicados, como el conocer el origen familiar y cultural. Sin embargo, se creyó que mediante el presente estudio se pudo dar una muy buena aproximación.

A continuación se presenta el análisis de cada uno de los clusters que fueron formados.

##### Interpretación Cluster 1

- ❖ **Tamaño del cluster** = 4 (0.8% del total de la muestra).
- ❖ **Edad promedio** = 21 – 30 (50% del clúster) 31 – 40 años (25% del clúster) y 41 – 100 años (25% del clúster).
- ❖ **Sexo** = Masculino (100% del clúster).
- ❖ **Estado civil** = Soltero (25% del clúster), Casado (25% del clúster), Divorciado (25% del clúster) y Conviviente (25% del clúster).
- ❖ **Nivel de instrucción** = Secundaria incompleta (5% del clúster) y Técnica (75% del clúster).
- ❖ **Reincidencia** = Reiterante (75% del clúster) y Reincidencia múltiple (25% del clúster).
- ❖ **Condición anímica** = Enfermo/Alterado (50% del clúster), Drogado (25% del clúster) y Ecuánime (25% del clúster).
- ❖ **Incidencia Horaria** = 8:01 a.m. - 2:00 p.m. (25% del clúster) 2:01 – 8:00 p.m. (50% del clúster) y 2:01 a.m. – 7:59 a.m. (25% del clúster).
- ❖ **Unidad de Intervención** = Comisaría El Porvenir (75% del clúster) y Comisaría La Victoria (25% del clúster).
- ❖ **Tipo de Arma** = Ninguna (50% del clúster) y Arma Blanca (50% del clúster).
- ❖ **Agresión** = Ambos tipos de agresión, es decir física y psicológica (100% del clúster).
- ❖ **Delito** = Delitos contra la Vida el Cuerpo y la Salud (100% de Lesiones).
- ❖ **Delito circunstancia** = Delito contra la Libertad (25% del clúster del delito Violación Libertad personal, y 50% del



clúster del delito Violación Domicilio) y Robo simple y agravado (25% del clúster).

- ❖ **Monto Robado** = No hubo robo (75% del clúster), Robo valorizado en más de 100 y menos de 500 soles (25% del clúster).

Se observó un clúster de varones, algunos han sido solteros y otros han tenido un compromiso como casados y convivientes, del cual el promedio de edad más predominante fue entre 21 años a más, con un nivel de instrucción predominante en el nivel técnico, aunque la distribución del nivel de instrucción del resto de la población del clúster fue más bajo aunque en proporciones pequeñas. En su mayoría fueron personas, que ya habían delinquido o han cometido el mismo delito, además al momento de cometer el delito estuvieron alteradas o bajo efectos de sustancias tóxicas como droga, aunque existieron otras personas que estuvieron conscientes de sus actos, es decir presentaron un estado de ecuanimidad.

Dichas personas delinquieron en contra de las personas, atentaron contra el delito de la Vida el Cuerpo y la Salud, específicamente cometieron el delito de Lesiones, en su mayoría han cometido este tipo de delitos en horas de la tarde-noche, en los alrededores de la ciudad de Chiclayo, ocasionando también delitos como delito de la libertad personal, es decir no solo han agredido a la persona sino que la han privado de su libertad y en caso de la violación de domicilio, es cuando han ingresado a una propiedad con el fin de herir a una persona en la mayoría de casos o por ocasionar algún tipo de robo, también algunas de estas personas no han poseído un arma y si han tenido solo ha sido un arma blanca, han utilizado la violencia tanto física y psicológica para cometer su delito, además los montos robados han sido menores oscilando entre los 100 y 500 soles, pues el objetivo principal no ha sido este sino el atentar contra la vida ajena.

Para que se cometiera un delito de lesiones o atentado contra la vida, en múltiples casos se da por ajustes de cuenta, por separación de relaciones o porque estas personas implicadas son contratadas por terceros para cometer este tipo de actos.

### Interpretación Clúster 2

- ❖ **Tamaño del cluster** = 6 (1.3% del total de la muestra).
- ❖ **Edad promedio** = 21 – 30 (83% del clúster) y 41 – 100 (17% del clúster).

- ❖ **Sexo** = Femenino (33% del clúster) y Masculino (67% del clúster).
- ❖ **Estado civil** = Casado (33% del clúster) y Conviviente (67% del clúster).
- ❖ **Nivel de instrucción** = Sin instrucción (17% del clúster), Primaria completa (33% del clúster), Secundaria incompleta (17% del clúster) y Secundaria completa (33% del clúster).
- ❖ **Reincidencia** = Primaria (100% del clúster).
- ❖ **Condición anímica** = Ebrio (50% del clúster) y Ecuánime (50% del clúster).
- ❖ **Incidencia Horaria** = 8:01 a.m – 2:00 p.m. (50% del clúster), 2:01 p.m. – 8:00 p.m. (17% del clúster) y 2:01 a.m. – 8:00 a.m. (33% del clúster).
- ❖ **Unidad de Intervención** = Campodónico (17% del clúster), Del Norte (67% del clúster) y Atusparia (17% del clúster).
- ❖ **Tipo de Arma** = Ninguna (83% del clúster) y Arma blanca (17% del clúster).
- ❖ **Agresión** = Física (17% del clúster), Psicológica (33% del clúster) y Ambas (50% del clúster).
- ❖ **Delito** = Delito contra la Familia (100% del clúster de Violencia familiar).
- ❖ **Delito circunstancia** = Robo simple y agravado (50% del clúster) y Daños simples y agravados (50% del clúster).
- ❖ **Monto Robado** = No hubo (16.7% del clúster), Robo valorizado hasta 100 soles (16.7% del clúster), Robo valorizado en más de 100 y menos o igual a 500 soles (16.7% del clúster), Robo valorizado en más de 500 y menos o igual a 1000 soles (33% del clúster) y Robo valorizado en más de 5000 (16.7% del clúster)

Se observó un clúster con personas en su mayoría varones (67%) y mujeres (33%), con un estado civil en mayor proporción convivientes y en menor proporción los casados, del cual el promedio de edad fue de 21 a más años, con un nivel de instrucción predominante en el nivel Primario y secundaria completa, aunque la distribución del nivel de instrucción del resto de la población del clúster fue un poco menor como el no tener ningún tipo de instrucción o no haber terminado un nivel como secundaria incompleta. En este tipo de delito por lo general la reincidencia es primaria, además al momento de cometer el delito mayormente estuvieron en un estado de ebriedad.

Dichas personas delinquieron contra la Familia, específicamente cometieron el delito de Violencia familiar, en su mayoría han cometido este tipo de delitos en horas de la

mañana, estas personas al momento de haber cometido el delito por lo general no han utilizado algún tipo de arma, es decir han utilizado la fuerza bruta, habiendo ocasionado mayormente agresiones tanto físicas como psicológicas, ambas a la vez, seguido de aquellos que han ocasionado agresiones físicas.

Para que se haya cometido un delito contra la familia debió ocurrir diferentes factores y múltiples circunstancias individuales. En el caso, el delito contra la familia fue más común que se cometiera por el varón, lo más probable fue que tratándose de personas con carga familiar, tengan problemas de incompatibilidad de caracteres, además se registró que los implicados tienen un nivel de educación en su mayoría medio-bajo y una reincidencia primaria, lo que significa que múltiples casos de agresiones o maltratos no son denunciados, y no porque se cometa por primera vez el delito sino porque la mayoría de personas atacadas son mujeres las cuales tienen miedo a denunciar dichos maltratos ya sea por imagen, miedo u otras causas, es por ello que la mayoría de implicados solo tienen una reincidencia primaria.

Asímimo la mayor frecuencia de estos delitos se ha dado cuando el implicado ha estado en estado de ebriedad y con arma blanca, la cual es la más común encontrar dentro de un Hogar (cuchillos, palos, etc.).

En principio, el clúster se identificó con las personas que cometen delitos en el ámbito familiar (mayormente los de violencia), asociados probablemente a factores como problemas de agresividad del implicado, crisis en el entorno familiar, problemas de alcoholismo, drogadicción, etc.

### **Interpretación Cluster 3**

- ❖ **Tamaño del cluster** = 13 (2.7% del total de la muestra).
- ❖ **Edad promedio** = 0 – 18 (7.7% del clúster), 19 – 20 (15.4% del clúster), 21 – 30 (15.4% del clúster), 31 – 40 (23.1% del clúster) y 41- 100 (38.5% del clúster).
- ❖ **Sexo** = Femenino (15.4% del clúster) y Masculino (84.6% del clúster).
- ❖ **Estado civil** = Soltero (30.8% del clúster), Casado (15.4% del clúster), Divorciado (7.7% del clúster), Conviviente (23.1% del clúster) y Exconviviente (23.1% del clúster).
- ❖ **Nivel de instrucción** = Primaria incompleta (46.2% del clúster), Primaria completa (15.4% del clúster),

Secundaria incompleta (23.1% del clúster) y Secundaria completa (15.4% del clúster).

- ❖ **Reincidencia** = Primaria (7.7% del clúster), Reiterante (46.2% del clúster) y Reincidencia Múltiple (46.2% del clúster),
- ❖ **Condición anímica** = Enfermo / Alterado (15.4% del clúster), Drogado (7.7% del clúster) y Ecuánime (76.9% del clúster).
- ❖ **Incidencia Horaria** = 8:01 a.m – 2:00 p.m. (23.1% del clúster), 2:01 p.m. – 8:00 p.m. (46.2% del clúster), 8:01 p.m. – 2:00 a.m. (15.4% del clúster) y 2:01 a.m. – 8:00 a.m. (15.4% del clúster)
- ❖ **Unidad de Intervención** = César Llatas Castro (15.4% del clúster), José Leonardo Ortiz (15.4% del clúster), El Porvenir (23.1% del clúster), Mujeres y Menores (15.4% del clúster) y Quiñones (30.8% del clúster).
- ❖ **Tipo de Arma** = Ninguna (61.5% del clúster), Arma blanca (15.4% del clúster), Arma de fuego (7.7% del clúster) y Ambas (15.4% del clúster).
- ❖ **Agresión** = No hubo (100% del clúster).
- ❖ **Delito** = Delitos contra el patrimonio (15.4% de Hurto simple y agravado y 84.6% de Robo simple y agravado del clúster).
- ❖ **Delito circunstancia** = Daños simples y agravados (100% del clúster).
- ❖ **Monto Robado** = Robo valorizado en más de 100 y menos o igual a 500 (30.8% del clúster), Robo valorizado más de 1000 y menos o igual a 5000 (15.4% del clúster) y Robo valorizado en más de 5000 (53.8% del clúster),

Este clúster agrupó a personas tanto mujeres como varones, sin embargo con mayor probabilidad en estos últimos, personas adultas entre 41 años a más, en su mayoría solteras, convivientes o exconvivientes, con un nivel de instrucción bajo (46.2% del nivel primaria incompleta) y con proporciones menores personas con primaria y secundaria completa.

Los implicados al momento de cometer un delito han estado en su mayoría conscientes (ecuanimes), en su mayoría tienen una reincidencia reiterante o reincidencia múltiple es decir que ya han cometido el mismo delito en otras ocasiones o han cometido otros delitos asociados a este como ocasionar daños no solo a las personas sino a su bienes materiales, los cuales han cometido un delito principalmente durante horas de la tarde-noche y por lo general no utilizan un arma al momento de delinquir.

Este clúster fue representado con delitos contra el Patrimonio, específicamente representado por Hurto simple y agravado y Robos simples y agravados, en los cuales no se cometió algún delito ya que estas personas han ingresado a propiedades ajenas para cometer sus actos delictivos aprovechando el descuido de los propietarios.

Además los montos robados con mayor frecuencia fueron mayores a 5000 soles.

Para el caso, lo más probable fue que tratándose de personas con un nivel de educación bajo y al no contar con los recursos para cubrir sus necesidades básicas, optan otras alternativas como fue el robar. Sin embargo no se pudo dejar de advertir, que el robo también se debió a otros factores como la exclusión del trabajo formal, la exclusión social, de salud, de educación, de vivienda, de participación social, problemas de alcoholismo, drogadicción sumado a ello como el haber crecido en medios hostiles y de violencia en que los delitos de robo no les eran ajeno. Estos factores, muchas veces se han traducido en una violación a la propiedad ajena.

#### **Interpretación Cluster 4**

- ❖ **Tamaño del cluster** = 137 (28.5% del total de la muestra).
- ❖ **Edad promedio** = 0 – 18 (6.6% del clúster), 19 – 20 (8% del clúster), 21 – 30 (47.4% del clúster), 31 – 40 (17.5% del clúster) y 41 – 100 (20.4% del clúster).
- ❖ **Sexo** = Femenino (11.7% del clúster) y Masculino (88.3% del clúster).
- ❖ **Estado civil** = Soltero (42.3% del clúster), Casado (10.9% del clúster), Divorciado (2.9% del clúster), Conviviente (34.3% del clúster), Exconviviente (8% del clúster) y Viudo (1.5% del clúster).
- ❖ **Nivel de instrucción** = Sin instrucción (5.8% del clúster), Primaria incompleta (11.7% del clúster), Primaria completa (18.2% del clúster), Secundaria incompleta (31.4% del clúster), Secundaria completa (29.9% del clúster) y Técnico (2.9% del clúster).
- ❖ **Reincidencia** = Primaria (12.4% del clúster), Reiterante (41.6% del clúster) y Reincidencia múltiple (46% del clúster).
- ❖ **Condición anímica** = Ebrio (3.6% del clúster), Enfermo / Alterado (6.6% del clúster), Drogado (5.8% del clúster) y Ecuánime (83.9% del clúster).
- ❖ **Incidencia Horaria** = 8:01 a.m – 2:00 p.m. (18.2% del clúster), 2:01 p.m. – 8:00 p.m. (43.8% del clúster), 8:01

p.m. – 2:00 a.m. (35% del clúster) y 2:01 – 8:00 a.m. (2.9% del clúster).

- ❖ **Unidad de Intervención** = César Llatas Castro (13.1% del clúster), JLO (8% del clúster), Campodónico (3.6% del clúster), Del Norte (2.2% del clúster), Atusparia (11.7% del clúster), El Porvenir (5.8% del clúster), Mujeres y Menores (16.8% del clúster), La Victoria (23.4% del clúster) y Quiñones (15.3% del clúster).
- ❖ **Tipo de Arma** = Ninguna (50.4% del clúster), Arma blanca (13.9% del clúster), Arma de fuego (32.1% del clúster) y Ambas (3.6% del clúster).
- ❖ **Agresión** = No hubo (70.8% del clúster), Física (29.2% del clúster).
- ❖ **Delito** = Delitos contra el patrimonio (20.4% de Hurto simple y agravado y 79.% de Robo simple y agravado del clúster).
- ❖ **Delito circunstancia** = No hubo (70.8% del clúster) y Lesiones (29.2% del clúster).
- ❖ **Monto Robado** = Robo valorizado hasta 100 soles (2.2% del clúster), Robo valorizado en más de 100 y menos o igual de 500 (29.2% del clúster), Robo valorizado en más de 500 y menos o igual a 1000 (14.6% del clúster), Robo valorizado en más de 1000 y menos de 5000 (38.7% del clúster) y más de 5000 (15.3% del clúster).

Este clúster agrupó a personas en su mayoría varones las cuales variaron entre jóvenes y adultos, con un nivel de educación bajo, los cuales en su mayoría solo alcanzaron a obtener un nivel primario o nivel secundario, además existen un porcentaje menor con personas que no estudiaron y un mínimo de personas con nivel técnico.

Las personas que fueron agrupadas dentro de este clúster presentaron una reincidencia múltiple en su mayoría sin embargo no se descarta que en otros casos existan personas reiterantes, los cuales cometieron un delito en mayor proporción durante horas de la tarde-noche.

El delito característico de este grupo fue contra el patrimonio, con mayor frecuencia en Robo simple y agravado, por ello fue común que no se registre agresiones en la mayoría de los casos, debido a que ese tipo de delito se comete rápidamente, aunque en proporciones menores si se ocasionaron agresiones físicas y se vincularon con otros delitos debido a la circunstancia del desarrollo del hecho, como son lesiones.

Finalmente los montos robados con mayor frecuencia oscilaron en cantidades variables entre menores a los 100 soles y aquellos

robos que con mayor proporción estuvieron por encima de los 1000 y 5000 soles.

### **Interpretación Cluster 5**

- ❖ **Tamaño del cluster** = 72 (15.9% del total de la muestra).
- ❖ **Edad promedio** = 0 – 18 (6.9% del clúster), 19 – 20 (5.6% del clúster), 21 – 30 (21% del clúster), 31 – 40 (32% del clúster) y 41 – 100 (35% del clúster).
- ❖ **Sexo** = Femenino (8.3% del clúster) y Masculino (92% del clúster).
- ❖ **Estado civil** = Soltero (25% del clúster), Casado (13% del clúster), Divorciado (4.2% del clúster), Conviviente (44% del clúster), Exconviviente (13% del clúster) y Viudo (1.4% del clúster).
- ❖ **Nivel de instrucción** = Sin instrucción (2.8% del clúster), Primaria incompleta (18% del clúster), Primaria completa (19% del clúster), Secundaria incompleta (31% del clúster), Secundaria completa (26% del clúster) y Técnico (2.8% del clúster).
- ❖ **Reincidencia** = Primaria (18% del clúster), Reiterante (56% del clúster) y Reincidencia múltiple (26% del clúster).
- ❖ **Condición anímica** = Ebrio (26% del clúster), Enfermo / Alterado (25% del clúster), Drogado (9.7% del clúster) y Ecuánime (39% del clúster).
- ❖ **Incidencia Horaria** = 8:01 a.m – 2:00 p.m. (36% del clúster), 2:01 p.m. – 8:00 p.m. (43% del clúster), 8:01 p.m. – 2:00 a.m. (17% del clúster) y 2:01 a.m. – 7:59 a.m. (4.2% del clúster).
- ❖ **Unidad de Intervención** = César Llatas (49% del clúster), JLO (35% del clúster), Campodónico (14% del clúster) y Del Norte (2.8% del clúster).
- ❖ **Tipo de Arma** = Ninguna (40% del clúster), Arma blanca (28% del clúster), Arma de fuego (31% del clúster) y Ambas (1.4% del clúster).
- ❖ **Agresión** = No hubo (74% del clúster) y Física (26% del clúster).
- ❖ **Delito** = Delito contra el Patrimonio (35% de Hurto simple y agravado y 65% de Robo simple y agravado del clúster).
- ❖ **Delito circunstancia** = No hubo (72% del clúster) y Lesiones (28% del clúster)
- ❖ **Monto Robado** = Robo valorizado hasta 100 soles (4.2% del clúster), Robo valorizado en más de 100 y menos o igual de 500 (29% del clúster), Robo valorizado en más de 500 y menos o igual a 1000 (39% del clúster), Robo

valorizado en más de 1000 y menos de 5000 (17% del clúster) y más de 5000 (11% del clúster).

Este clúster agrupo a personas en su mayoría varones de entre 21 a más años, con mayor frecuencia convivientes, seguido de solteros y en porcentajes menores los casados y exconvivientes, con un nivel de instrucción variado, desde los de sin instrucción hasta los del nivel técnico, sin embargo la mayor concentración fue en los de nivel secundaria incompleta el cual ocupó el 31%.

Los implicados al momento de cometer un delito han estado mayormente en estado de ecuanimidad y principalmente han sido reiterantes, es decir ya han cometido el mismo delito, esto se refiere a delitos contra el patrimonio y en muchas ocasiones a pesar de las denuncias impuestas estas personas siguen cometiendo el mismo delito, quienes cometieron un delito principalmente durante horas de la tarde-noche, en su mayoría los implicados ya cuentan con ubicaciones de robo y por haber cometido un mismo delito una y otra vez por lo general no utilizan algún tipo de arma. Aunque existe un porcentaje menor de implicados que si utilizan un tipo de arma como fue arma blanca para cometer el delito.

Este clúster fue representado contra delitos contra el Patrimonio, específicamente Hurto y Robo simple y agravado, es por ello que son bajas las posibilidades de agresión, por lo general son suscitadas cuando la víctima muestra resistencia.

Así mismo los montos robados oscilan desde menos o igual a 100 soles generalmente cuando se produce un hurto, como es robo de una cartera, un celular, etc. hasta robos que oscilan en más de 5000 soles.

#### **Interpretación Cluster 6**

- ❖ **Tamaño del cluster** = 99 (20.6% del total de la muestra).
- ❖ **Edad promedio** = 19 – 20 (2% del clúster), 21 – 30 (33.3% del clúster), 31 – 40 (30.3% del clúster) y 41 – 100 (34.3% del clúster).
- ❖ **Sexo** = Femenino (36.4% del clúster) y Masculino (63.6% del clúster).
- ❖ **Estado civil** = Soltero (5% del clúster), Casado (40.4% del clúster), Conviviente (45.5% del clúster), Exconviviente (8% del clúster) y Viudo (1% del clúster).
- ❖ **Nivel de instrucción** = Primaria incompleta (3% del clúster), Primaria completa (24.2% del clúster), Secundaria incompleta (20.2% del clúster), Secundaria



completa (29.3% del clúster) , Técnico (9% del clúster) y Superior (14.1% del clúster).

- ❖ **Reincidencia** = Primaria (89.9% del clúster) y Reiterante (10.1% del clúster).
- ❖ **Condición anímica** = Ebrio (23.2% del clúster), Drogado (1% del clúster) y Ecuánime (75.8% del clúster).
- ❖ **Incidencia Horaria** = 8:00 a.m – 2:00 p.m. (39.4% del clúster), 2:01 p.m. – 8:00 p.m. (39.4% del clúster), 8:01 p.m. – 2:00 a.m. (19.2% del clúster) y 2:01 a.m. – 8:00 a.m. (2% del clúster).
- ❖ **Unidad de Intervención** = César Llatas (5.1% del clúster), JLO (11.1% del clúster), Campodónico (5.1% del clúster), Del Norte (53.5% del clúster), Atusparia (8.1% del clúster), El Porvenir (7.1% del clúster), Mujeres y Menores (3% del clúster), La Victoria (4% del clúster) y Quiñones (3% del clúster).
- ❖ **Tipo de Arma** = Ninguna (93.9% del clúster) y Arma blanca (6.1% del clúster).
- ❖ **Agresión** = No hubo (2% del clúster), Física (24.2% del clúster), Psicológica (43.4% del clúster) y Ambas (30.3% del clúster).
- ❖ **Delito** = Delito contra la familia (100% de Violencia familiar del clúster).
- ❖ **Delito circunstancia** = No hubo (100% del clúster).
- ❖ **Monto Robado** = No hubo (100% del clúster).

Se observó un clúster con personas en su mayoría varones (63.6%) y mujeres (36.4%), con un estado civil predominante en los convivientes y casados, con un nivel de instrucción predominante en el nivel Secundaria completa, aunque la distribución del nivel de instrucción del resto de la población del clúster fue variada con un nivel más bajo como lo fueron las personas de nivel primaria incompleta y también con un nivel más elevado siendo este un nivel Superior (14.1%). En su mayoría fueron personas, que por primera vez habían sido denunciadas, además al momento de cometer el delito un 75.8% de estas personas estuvieron en un estado de ecuanimidad y en proporciones menores personas en estado de ebriedad.

Dichas personas delinquieron contra la Familia, específicamente cometieron el delito de Violencia familiar, en su mayoría han cometido este tipo de delitos en horas de la mañana-tarde, estas personas al momento de haber cometido el delito por lo general no han utilizado ningún arma en la mayoría de los casos, sin embargo en otros casos se ha utilizado un arma blanca, esto resulta más común debido a que son

delitos del entorno familiar y este tipo de armas se puede encontrar fácilmente en un hogar como piedras, cuchillos, etc.

Principalmente cuando se comete este tipo de delitos no solo se agrede físicamente sino psicológicamente, es decir ambos tipos de agresiones a la vez.

De los registros se supo que cuando suceden este tipo de hechos por lo general no se busca cometer otro delito o sustraer algún bien material, en múltiples casos este tipo de agresiones se debió a problemas anteriormente ocurridos dentro del ámbito familiar y no específicamente al problema que dio origen la denuncia, en este tipo de delitos es muy común que no se denuncie cuando se detecta por primera vez el problema o inclusive no se llegue a denunciar, es por ello que cuando se denuncia la gran mayoría de las víctimas manifiestan que no era la primera vez que sucede ese tipo de agresiones.

#### **Interpretación Cluster 7**

- ❖ **Tamaño del cluster** = 12 (2.5% del total de la muestra).
- ❖ **Edad promedio** = 21 – 30 (16.7% del clúster), 31 – 40 (16.7% del clúster) y 41 – 100 (66.7% del clúster).
- ❖ **Sexo** = Femenino (33.3% del clúster) y Masculino (66.7% del clúster).
- ❖ **Estado civil** = Casado (8.3% del clúster), Conviviente (66.7% del clúster) y Exconviviente (25% del clúster).
- ❖ **Nivel de instrucción** = Sin instrucción (16.7% del clúster), Primaria incompleta (16.7% del clúster), Primaria completa (16.7% del clúster), Secundaria incompleta (8.3% del clúster) y Secundaria completa (41.7% del clúster).
- ❖ **Reincidencia** = Primaria (8.3% del clúster), Reiterante (41.7% del clúster) y Reincidencia múltiple (50% del clúster).
- ❖ **Condición anímica** = Ebrio (16.7% del clúster), Enfermo / Alterado (25% del clúster), Drogado (25% del clúster) y Ecuánime (33.3% del clúster).
- ❖ **Incidencia Horaria** = 8:01 a.m – 2:00 p.m. (25% del clúster), 2:01 p.m. – 8:00 p.m. (25% del clúster), 8:01 p.m. – 2:00 a.m. (41.7% del clúster) y 2:01 a.m. – 8:00 a.m. (8.3% del clúster).
- ❖ **Unidad de Intervención** = César Llatas Castro (16.7% del clúster), Campodónico (25% del clúster), Del Norte (16.7% del clúster), Atusparia (16.7% del clúster) Mujeres y Menores (16.7% del clúster) y La Victoria (8.3% del clúster).

- ❖ **Tipo de Arma** = Arma blanca (25% del clúster), Arma de fuego (58.3% del clúster) y Ambas (16.7% del clúster).
- ❖ **Agresión** = Física (25% del clúster) y Psicológica (75% del clúster).
- ❖ **Delito** = Delito contra la Vida el Cuerpo y la Salud (100% de Lesiones).
- ❖ **Delito circunstancia** = Delito contra la Libertad (83.3% de Violación libertad personal y Violación domicilio 16.7% del clúster).
- ❖ **Monto Robado** = No hubo robo (100% del clúster).

Este clúster agrupó a personas del sexo masculino de más de 21 años, convivientes en la mayoría de casos, quienes han cometido delitos pero no por primera vez sino de forma múltiple en su gran mayoría, se refiere a personas con un alto historial criminalístico.

Las personas implicadas al momento de haber cometido el delito en gran proporción han estado totalmente conscientes (sobrias), y por lo general han cometido este tipo de actos durante la noche. Además este tipo de personas para intimidar a su víctima y cometer su delito utiliza en gran parte un tipo de arma de fuego.

Este clúster estuvo representado por delitos contra la vida el cuerpo y la salud, fue por ello que el porcentaje de ocasión de agresiones son altas pues se busca atentar contra la vida ajena, sin embargo existieron otros delitos circunstanciales que se cometieron producto de los delitos contra la vida el cuerpo y la salud en su mayoría estuvo representado por delitos contra la libertad personal y en proporciones menores de los de violación de domicilio, en los cuales se vio afectada la integridad de la persona y/o de sus bienes materiales.

Este tipo de delitos frecuentemente es ocasionado por venganza, ajuste de cuentas, problemas entre vecinos, celos etc., en los cuales por lo general se generó una agresión tipo psicológica, es decir como también se desarrollaron otro tipo de delitos como fueron el atentar contra la libertad de la persona, lo que se busca es minimizar al agraviado.

Generalmente este tipo de delito únicamente tiene como finalidad atentar contra el bienestar de la persona, por diversos motivos como los mencionados anteriormente por ello no se comete algún tipo de apropiación de bienes ajenos.

### **Interpretación Cluster 8**

- ❖ **Tamaño del cluster** = 8 (1.7% del total de la muestra).
- ❖ **Edad promedio** = 21 – 30 (25% del clúster), 31 – 40 (25% del clúster) y 41 – 100 (50% del clúster).
- ❖ **Sexo** = Femenino (75% del clúster) y Masculino (25% del clúster).
- ❖ **Estado civil** = Soltero (13% del clúster), Conviviente (13% del clúster) y Exconviviente (75% del clúster).
- ❖ **Nivel de instrucción** = Sin instrucción (13% del clúster), Primaria incompleta (13% del clúster), Primaria completa (13% del clúster), Secundaria incompleta (25% del clúster), Secundaria completa (25% del clúster) y Técnico (13% del clúster).
- ❖ **Reincidencia** = Primaria (38% del clúster), Reiterante (50% del clúster) y Reincidencia múltiple (13% del clúster).
- ❖ **Condición anímica** = Ebrio (25% del clúster), Enfermo / Alterado (38% del clúster), Drogado (25% del clúster) y Ecuánime (13% del clúster).
- ❖ **Incidencia Horaria** = 8:01 a.m – 2:00 p.m. (50% del clúster), 2:01 p.m. – 8:00 p.m. (38% del clúster) y 2:01 a.m. – 8:00 a.m. (13% del clúster).
- ❖ **Unidad de Intervención** = JLO(25% del clúster), El Porvenir (37.5% del clúster), Mujeres y Menores (25% del clúster), y Quiñones (12.5% del clúster).
- ❖ **Tipo de Arma** = Ninguna (12.5% del clúster), Arma blanca (12.5% del clúster), Arma de fuego (62.5% del clúster) y Ambas (12.5% del clúster).
- ❖ **Agresión** = Ambas (100% del clúster).
- ❖ **Delito** = Delito contra la Vida el Cuerpo y la Salud (100% de Lesiones del clúster).
- ❖ **Delito circunstancia** = Delito contra la Humanidad (100% de delito Desaparición Forzada del clúster).
- ❖ **Monto Robado** = No hubo robo (100% del clúster).

Se observó un clúster predominante en el sexo femenino, con estado civil más notorio fue el ser separada (exconviviente), seguido de las personas convivientes y casadas en proporciones menores, quienes han tenido un nivel de instrucción en mayor frecuencia en el nivel Secundario, aunque la distribución del nivel de instrucción del resto de la población del clúster fue muy variable siendo el nivel más bajo las personas sin instrucción y esporádicamente los de nivel técnico. En su mayoría fueron personas que han sido reiterantes, es decir ya han cometido más de una vez el mismo delito, además al momento de cometer el delito un gran porcentaje ha estado alterado o bajo efecto de alguna sustancia nociva.

Dichas personas delinquieron contra la Vida el Cuerpo y la Salud, específicamente cometieron el delito de Lesiones, en su mayoría han cometido este tipo de delitos en horas de la mañana, estas personas al momento de haber cometido el delito especialmente han utilizado un arma de fuego por lo general han ocasionado agresiones a su víctima tanto física como psicológicas.

De los registros se supo que cuando suceden este tipo de hechos por lo general conlleva a cometer otro delito, como la desaparición forzada de la víctima de forma momentánea.

### **Interpretación Cluster 9**

- ❖ **Tamaño del cluster** = 70 (14.6% del total de la muestra).
- ❖ **Edad promedio** = 19 – 20 (2.9% del clúster), 21 – 30 (10% del clúster), 31 – 40 (30% del clúster) y 41 – 100 (57% del clúster).
- ❖ **Sexo** = Femenino (43% del clúster) y Masculino (57% del clúster).
- ❖ **Estado civil** = Soltero (10% del clúster), Casado (21% del clúster), Divorciado (7.1% del clúster), Conviviente (26% del clúster), Exconviviente (31% del clúster) y Viudo (4.3% del clúster).
- ❖ **Nivel de instrucción** = Sin instrucción (7.1% del clúster), Primaria incompleta (8.6% del clúster), Primaria completa (11% del clúster), Secundaria incompleta (20% del clúster), Secundaria completa (37% del clúster) y Técnica (16% del clúster).
- ❖ **Reincidencia** = Primaria (87% del clúster) y Reiterante (13% del clúster).
- ❖ **Condición anímica** = Ebrio (27% del clúster), Enfermo/Alterado (23% del clúster), Drogado (21% del clúster) y Ecuánime (29% del clúster).
- ❖ **Incidencia Horaria** = 8:01 a.m – 2:00 p.m. (21% del clúster), 2:01 p.m. – 8:00 p.m. (19% del clúster), 8:01 p.m. – 2:00 a.m. (34% del clúster) y 2:01 a.m. – 8:00 a.m. (26% del clúster).
- ❖ **Unidad de Intervención** = César Llatas Castro (8.6% del clúster), JLO (4.3% del clúster), Campodónico (1.4% del clúster), Del Norte (1.4% del clúster), Atusparia (5.7% del clúster), El Porvenir (20% del clúster), Mujeres y Menores (27% del clúster), La Victoria (17% del clúster) y Quiñones (14% del clúster).
- ❖ **Tipo de Arma** = Ninguna (57% del clúster), Arma blanca (31% del clúster), Arma de fuego (10% del clúster) y Ambas (1.4% del clúster).

- ❖ **Agresión** = Física (77% del clúster), Psicológica (5.7% del clúster) y Ambas (17% del clúster).
- ❖ **Delito** = Delito contra la Vida el Cuerpo y la Salud (100% de Lesiones).
- ❖ **Delito circunstancia** = No hubo (100% del clúster).
- ❖ **Monto Robado** = No hubo (100% del clúster).

Se observó un clúster de personas en su mayoría fueron varones, con un mayor porcentaje exconvictos, entre jóvenes y adultos, con un nivel de instrucción predominante en el nivel secundario, aunque la distribución del nivel de instrucción del resto de la población del clúster fue muy variable con un nivel menor se encuentran las personas sin instrucción y de nivel primario, y con un porcentaje más elevado a personas de nivel técnico. En su mayoría fueron personas, que por primera vez estaban implicadas en un delito, sin embargo con un porcentaje menor las personas reiterantes, quienes al momento de cometer el delito estuvieron en su mayoría en un estado de ecuanimidad.

Dichas personas delinquieron contra la Vida el cuerpo y la Salud, específicamente cometieron el delito de Lesiones, en su mayoría este tipo de delito han sido cometidos en horas de la noche, no solo en los alrededores de la ciudad de Chiclayo sino también en el centro de la ciudad, habiendo ocasionado mayormente agresiones físicas, sin embargo le siguió la agresión psicológica y física que fueron muy comunes para reducir a su víctima, este tipo de delitos cuando se trata de conflictos con personas relativamente cercanas por lo general no requieren de un arma sin embargo en otros casos utilizan un arma blanca.

El clúster refleja que no se ha buscado provocar otro tipo de delito adicional y lo más probable fue que tratándose de personas con un nivel de educación medio-bajo, cuando se les presenta un problema en su entorno fácilmente lo resuelven con insultos, golpes, etc., convirtiendo de esa forma a la violencia en un símbolo de respeto, en una forma de manifestar la presión.

#### **Interpretación Cluster 10**

- ❖ **Tamaño del cluster** = 32 (6.7% del total de la muestra).
- ❖ **Edad promedio** = 18 – 20 (9.4% del clúster), 21 – 30 (28.1% del clúster), 31 – 40 (3.1% del clúster) y 41 – 100 (59.4% del clúster).

- ❖ **Sexo** = Femenino (40.6% del clúster) y Masculino (59.4% del clúster).
- ❖ **Estado civil** = Soltero (34.4% del clúster), Casado (18.8% del clúster), Divorciado (3.1% del clúster), Conviviente (6.3% del clúster), Exconviviente (34.4% del clúster) y Viudo (3.1% del clúster).
- ❖ **Nivel de instrucción** = Sin instrucción (6.3% del clúster), Primaria incompleta (15.6% del clúster), Primaria completa (21.9% del clúster), Secundaria incompleta (15.6% del clúster), Secundaria completa (21.9% del clúster) y Técnica (18.8% del clúster).
- ❖ **Reincidencia** = Primaria (9.4% del clúster), Reiterante (62.5% del clúster) y Reincidencia múltiple (28.1% del clúster).
- ❖ **Condición anímica** = Ebrio (34.4% del clúster), Enfermo / Alterado (25% del clúster), Drogado (21.9% del clúster) y Ecuánime (18.8% del clúster).
- ❖ **Incidencia Horaria** = 8:01 a.m – 2:00 p.m. (15.6% del clúster), 2:01 p.m. – 8:00 p.m. (50% del clúster), 8:01 p.m. – 2:00 a.m. (18.8% del clúster) y 2:01 a.m. – 8:00 a.m. (15.6% del clúster)
- ❖ **Unidad de Intervención** = César Llatas Castro (12.5% del clúster), JLO (21.9% del clúster), Campodónico (31.3% del clúster), Del Norte (9.4% del clúster), Atusparia (15.6% del clúster), Mujeres y Menores (3.1% del clúster), La Victoria (3.1% del clúster) y Quiñones (3.1% del clúster).
- ❖ **Tipo de Arma** = Ninguna (84.4% del clúster) , Arma blanca (3.1% del clúster), Arma de fuego (3.1% del clúster) y Ambas (9.4% del clúster).
- ❖ **Agresión** = Física (62.5% del clúster), Psicológica (6.3% del clúster) y Ambas (31.3% del clúster).
- ❖ **Delito** = Delito contra la Vida el Cuerpo y la Salud (100% de Lesiones del clúster).
- ❖ **Delito circunstancia** = No hubo (100% del clúster).
- ❖ **Monto Robado** = No hubo (100% del clúster).

Se observó un clúster en su mayoría varones, entre jóvenes y adultos con mayor frecuencia aquellas personas sin compromiso, es decir solteros seguido por los exconvivientes, con un nivel de instrucción predominante en el nivel secundario, aunque la distribución del nivel de instrucción del resto de la población del clúster fue muy bajo siendo personas sin instrucción o del nivel primario, y minimamente han logrado un nivel técnico. En su mayoría fueron personas, que han sido reiterantes, anteriormente ya han estado implicadas en un delito, además al momento de cometer el delito han estado bajo efectos de sustancias tóxicas- nocivas como es el alcohol o drogas.

Dichas personas delinquieron contra la Vida el Cuerpo y la Salud, específicamente cometieron el delito de Lesiones, en su mayoría han cometido este tipo de delitos en horas de la tarde-noche.

Existen múltiples variables pero de los datos registrados en gran medida se debe a efectos nocivos en las personas con lo cual afectaron la integridad de la víctima al ocasionarle maltratos físicos o psicológicos.

En algunos de los casos no se contó con un arma en su mayoría, pero existieron casos en los que si utilizaron armas de fuego o armas blancas.

En el caso, lo más probable fue que tratándose de personas que al estar bajo la presencia de sustancias tóxicas busquen violentar a otras personas ya sea por placer o por necesidad de conseguir dinero para sus vicios, pues la gran mayoría de estas personas están desempleadas.



### Interpretación Cluster 11

- ❖ **Tamaño del cluster** = 6 (1.3% del total de la muestra).
- ❖ **Edad promedio** = 18 – 20 (33.3% del clúster), 21 – 30 (50% del clúster) y 31 – 40 (16.7% del clúster).
- ❖ **Sexo** = Femenino (33.3% del clúster) y Masculino (66.7% del clúster).
- ❖ **Estado civil** = Soltero (66.7% del clúster), Casado (16.7% del clúster) y Conviviente (16.7% del clúster).
- ❖ **Nivel de instrucción** = Secundaria incompleta (66.7% del clúster) y Secundaria completa (33.3% del clúster).
- ❖ **Reincidencia** = Reiterante (100% del clúster).
- ❖ **Condición anímica** = Ecuánime (100% del clúster).
- ❖ **Incidencia Horaria** 2:01 a.m. – 8:00 a.m. (100% del clúster).
- ❖ **Unidad de Intervención** = El Porvenir (100% del clúster).
- ❖ **Tipo de Arma** = Arma de fuego (100% del clúster).
- ❖ **Agresión** = Ambas (100% del clúster).
- ❖ **Delito** = Delito contra el Patrimonio (100% de Usurpación / Extorsión del clúster).
- ❖ **Delito circunstancia** = Delito contra la Libertad (100% de Violación Libertad Personal del clúster).
- ❖ **Monto Robado** = Robo valorizado en más de 5000 (100% del clúster).

Se observó un clúster en su mayoría varones, sin embargo en proporciones menores mujeres, en el cual predominó como estado civil el ser soltero, los cuales fueron en su mayoría jóvenes, con un nivel de instrucción predominante en el nivel secundaria incompleta. Estas personas son expertas en extorsión y ya han estado implicados en casos anteriores, además al momento de cometer el delito estuvieron en estado de ecuanimidad para lograr llevar a cabo sus planes.

Dichas personas delinquieron contra el Patrimonio, específicamente cometieron el delito de Usurpación/ Extorsión, en su mayoría han cometido este tipo de delitos en horas de la madrugada.

Por lo general este tipo de delitos, busca primero obtener información de su víctima, ya sea por medio del trabajo o por una persona cercana al agraviado, además para llevar a cabo este tipo de delitos los delincuentes utilizaron arma de fuego, con lo cual genera mayor presión en su víctima y lograr robarle un monto alto, es decir según el clúster los montos robados oscilan en más de 5000 soles.

## Interpretación Cluster 12

- ❖ **Tamaño del cluster** = 21 (4.4% del total de la muestra).
- ❖ **Edad promedio** = 21 – 30 (38.1% del clúster), 31 – 40 (28.6% del clúster) y 41 – 100 (33.3% del clúster).
- ❖ **Sexo** = Femenino (9.5% del clúster) y Masculino (90.5% del clúster).
- ❖ **Estado civil** = Soltero (4.8 % del clúster), Casado (9.5% del clúster), Conviviente (76.2% del clúster), Exconviviente (4.8% del clúster) y Viudo (4.8% del clúster).
- ❖ **Nivel de instrucción** = Sin Instrucción (9.5% del clúster), Primaria incompleta (9.5% del clúster), Primaria completa (23.8% del clúster), Secundaria incompleta (38.1% del clúster), Secundaria completa (14.3% del clúster) y Técnico (4.8% del clúster).
- ❖ **Reincidencia** = Primaria (4.8% del clúster), Reiterante (61.9% del clúster) y Reincidencia múltiple (33.3% del clúster).
- ❖ **Condición anímica** = Ebrio (4.8% del clúster), Enfermo / Alterado (4.8% del clúster), Ecuánime (90.5% del clúster).
- ❖ **Incidencia Horaria** = 2:01 p.m. – 8:00 p.m. (38.1% del clúster) y 8:01 p.m. – 2:00 a.m. (61.9% del clúster).
- ❖ **Unidad de Intervención** = César Llatas (76.2% del clúster), Campodónico (9.5% del clúster) y El Porvenir (14.3% del clúster),
- ❖ **Tipo de Arma** = Ninguna (38.1% del clúster), Arma blanca (14.3% del clúster) y Arma de fuego (47.7% del clúster).
- ❖ **Agresión** = Ninguna (33.3% del clúster), Física (14.3% del clúster), Psicológica (33.3% del clúster) y Ambas (19% del clúster).
- ❖ **Delito** = Delito contra el Patrimonio (19.1% de Daños simples y agravados y 80.9 Usurpación / Extorsión del clúster).
- ❖ **Delito circunstancia** = No hubo (66.7% del clúster), Delito contra la Vida el Cuerpo y La Salud (33.3% Lesiones del clúster).
- ❖ **Monto Robado** = Robo valorizado en más de 1000 y menos o igual d 5000 (28.6% del clúster) y Robo valorizado en más de 5000 (71.4% del clúster).

Se observó un clúster en su mayoría varones, sin embargo en proporciones menores mujeres, en el cual predominó como estado civil el de convivencia, del cual han sido entre jóvenes y adultos, con un nivel de instrucción predominante en el nivel secundario, aunque la distribución del nivel de instrucción del resto de la población del clúster fue variada, algunos tuvieron

un nivel de instrucción menor como personas sin instrucción o del nivel primario, y en otros alcanzaron un nivel técnico. En su mayoría fueron personas reiterantes en un delito, además al momento de cometer el delito la mayoría estuvieron en un estado de ecuanimidad.

Dichas personas delinquieron contra el patrimonio, específicamente cometieron el delito de Usurpación / Extorsión en mayor cantidad y en proporciones menores el delito de Daños simples y agravados, en su mayoría han cometido este tipo de delitos en horas de la noche.

Por lo general los delitos de usurpación / extorsión fue común que se tuviera un vínculo indirecto a la persona agraviada o para el caso del delito de daños simples y agravados que se busca obtener dinero de los robos ingresando a las propiedades ajenas sin presencia del agraviado. Para tales casos el fin no es afectar la integridad de las personas sino obtener bienes materiales o monetarios.

#### **4.3.3.2.2. Clasificación**

Después de la obtención de los grupos se procedió a realizar la clasificación supervisada, para lo cual se utilizó Redes Neuronales haciendo uso del Matlab R2010a.

Para lo cual se definieron las capas de entradas, ocultas y de salidas, se utilizaron 13 entradas correspondientes a las variables definidas en la etapa de clustering, para conocer el número óptimo de capas ocultas, se contrastó el número de entradas con las salidas, dando como resultado una red oculta, para optimizar el trabajo, lo que evitó el sobreajuste dentro del entrenamiento, y para la capas de salida son 12 referidos a los grupos ya establecidos.

Se utilizaron 150 iteraciones, fue recomendable no dar un número reducido ya que la red no puede lograr el propósito de entrenamiento.

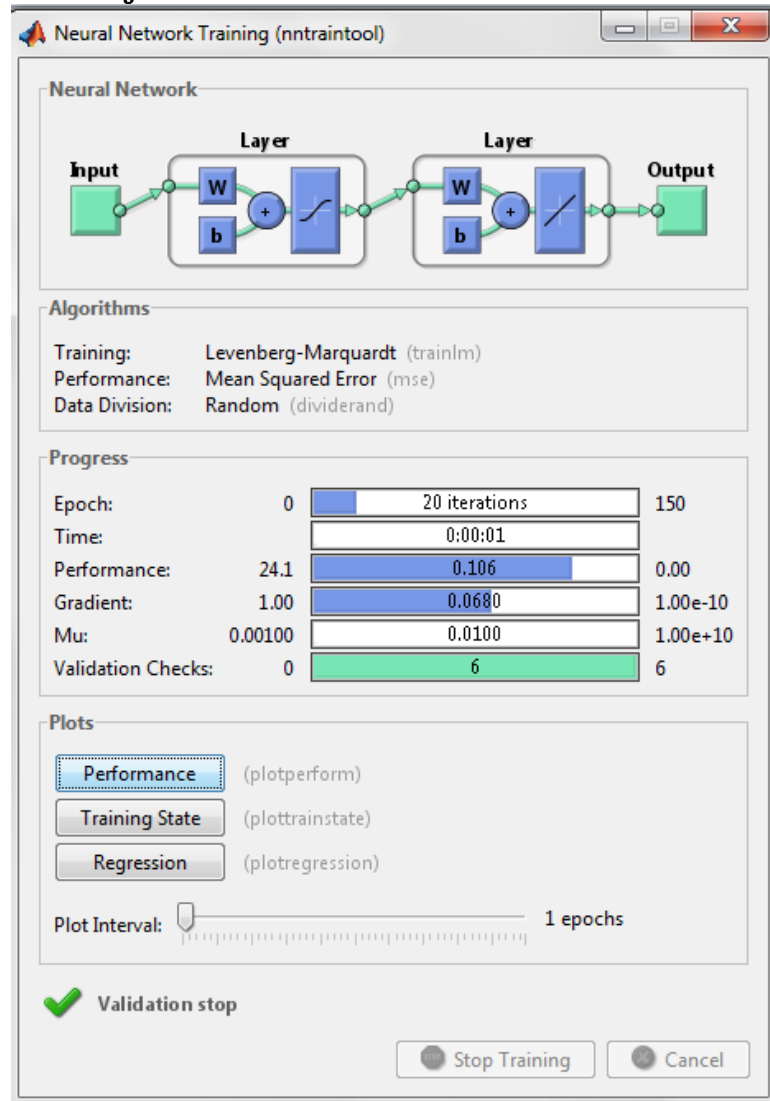
Además para el proceso de la red neuronal, se consideró tres fases:

Fase de entrenamiento, en la cual se utilizaron las 359 denuncias válidas para determinar los parámetros que definen el modelo de red neuronal.

Después pasó a la etapa de validación, esta etapa se utilizó para evitar el sobreajuste, es decir en esta etapa permitió controlar el proceso de aprendizaje.

Finamente se encuentra la etapa de testeo, conformado por un pequeño grupo de datos que permitió verificar el correcto funcionamiento.

**Fig. 18: Entrenamiento de la Red Neuronal**



#### 4.3.4. Fase V: Evaluación

##### 4.3.4.1. Evaluación de Resultados

La evaluación desde todo punto de vista resultó ser muy positiva. Se comenzó el proyecto desconociendo totalmente el tema en cuestión y mediante las técnicas de minería de datos se pudo llegar a conclusiones interesantes, de acuerdo a la opinión vertida por los expertos del dominio.

Otro punto a favor es que la interpretación de los resultados obtenidos a través de los algoritmos utilizados, tanto para el

agrupamiento de los datos como para la clasificación, se encontró alineada con la opinión de los especialistas en materia delictiva.

Los especialistas en materia delictiva coincidieron en un punto clave:

“Para lograr un mejor entendimiento de las conductas y comportamientos del implicado, es sumamente importante conocer su historial, familiar y socio-cultural del mismo”.

Lamentablemente, no se contó con dicha información en los datos de sus registros y su historial delictivo estuvo desligado de un caso a otro.

Conclusiones de las principales reglas de decisión a partir de la aplicación de los algoritmos y conocimientos adquiridos:

La mayoría de las personas implicadas en un principio fueron por haber cometido un delito contra el patrimonio específicamente el delito de Hurto simple y agravado y Robo simple y agravado seguido por el delito en contra de la vida el cuerpo y la salud.

Existe una distribución del nivel de instrucción muy pobre entre los implicados.

Aun cuando una persona sea soltera o tenga un nivel de instrucción regular, comete delitos contra la propiedad. Pudo estar relacionado con la cultura del robo que trae innata.

La mayoría de las personas implicadas que delinquen contra la familia específicamente violencia familiar son varones y tienen una reincidencia primaria. Esto pudo ser ligado a que la mayoría de personas agraviadas son mujeres y se encontraron en una situación de sometimiento.

El nivel de instrucción estuvo fuertemente relacionado con el delito de robo y hurto en personas menores a 31 años.

Los delitos de violencia familiar y lesiones a las personas en múltiples ocasiones se han dado porque el implicado ha estado bajo efectos de sustancias nocivas como alcohol, drogas, etc., o alterado a al momento de cometer el delito.

En general, quienes cometen delitos contra la integridad de la persona, son personas en su mayoría exconvictos, mayores a 31 años y con mayor probabilidad de ocurrencia del delito durante horas de la madrugada.

#### 4.4. FASE 03: CONSTRUCCIÓN

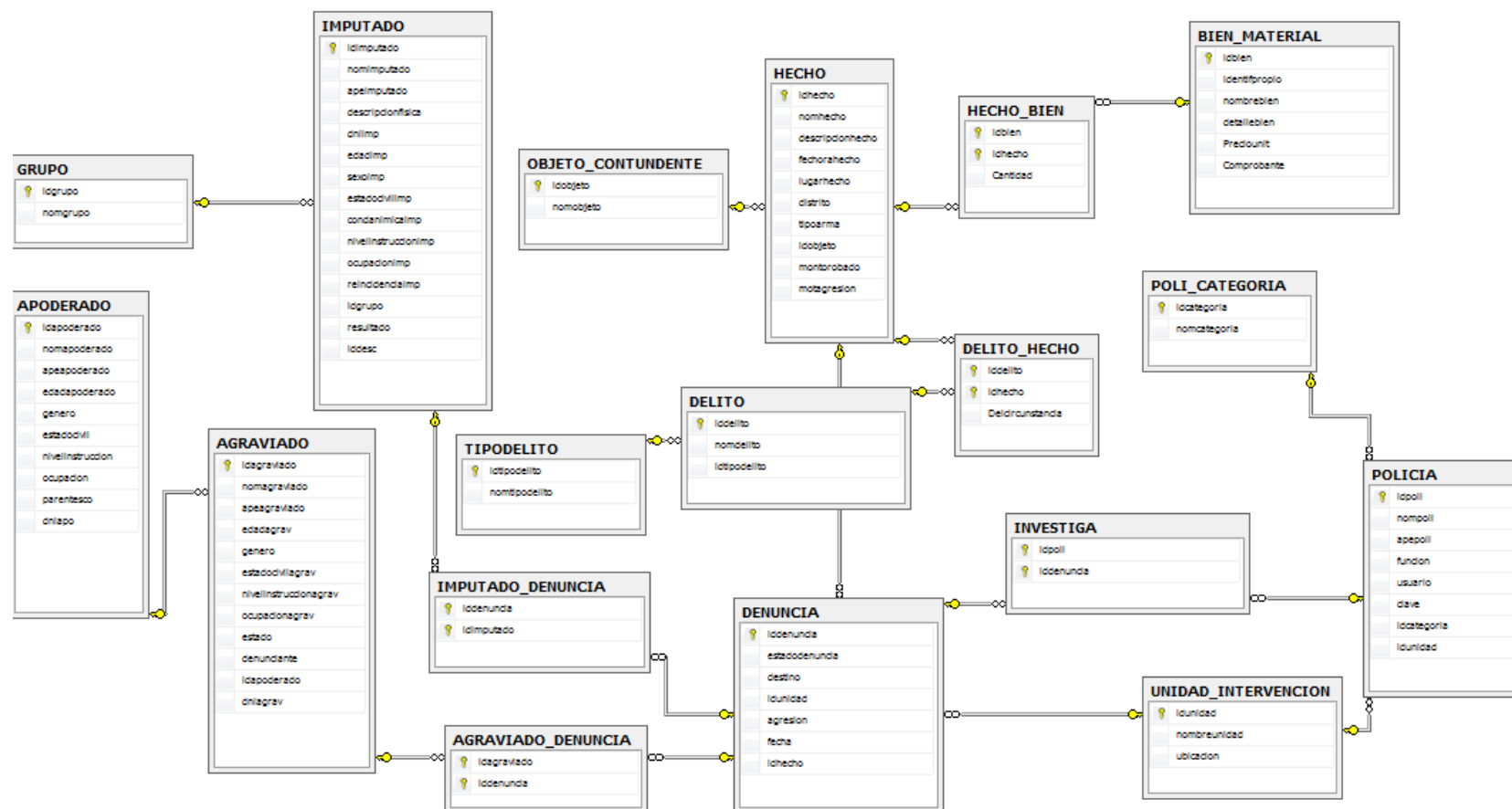
**Desarrolle iterativamente un producto completo que esté listo para hacer transición a su comunidad de usuarios:** Se desarrollaron los requisitos restantes, se completaron en detalles los diseños, se completó la implementación y prueba del software. Se liberó la primera versión operativa del software (beta), del Sistema de Alerta Temprana y se determinó que los policías estaban listos para que la aplicación sea desplegada.

**Diseñar, implementar, validar y establecer la línea base para la arquitectura:**

Se diseñó el modelo de Base de Datos el cual dio soporte al Sistema de Alerta Temprana. (Figura 19)

Se diseñó, implementó y probó un esqueleto estructural del Sistema de Alerta Temprana. Aunque la funcionalidad no estuvo completa aún, muchas de las interfaces entre los bloques de construcción fueron implementadas y probadas.

Fig. 19. Base de Datos del Sistema



#### 4.5. FASE 04: TRANSICIÓN

Se aseguró que el Sistema de Alerta Temprana está listo para ser distribuido a los usuarios y se llevó a cabo su implementación.

##### Descripción de la Interfaces

La primera interfaz permitió al usuario acceder a las funcionalidades presentes en el Sistema de Alerta Temprana, se definieron 2 **Categorías de Usuario** (Administrador, Subalterno para cada uno de ellos se determinaron un conjunto de funcionalidades), luego se ingresa el nombre de Usuario, y a continuación introducir la contraseña.

En esta interfaz se puede dar fe que se cumple con la **ETAPA DE ACCESO AL SISTEMA** ya que por medio de esta se podrá acceder a todo el proceso de Registro de denuncias y definición de perfiles delictivos.

Fig. 20. Acceso al Sistema



Una vez que se han ingresado los campos requeridos, se da clic en el botón **Ingresar** y si los datos son correctos a continuación se nos presentará un mensaje de bienvenida (Figura 21) caso contrario se nos mostrará un mensaje de Error (Figura 22).

Fig. 21: Mensaje de bienvenida

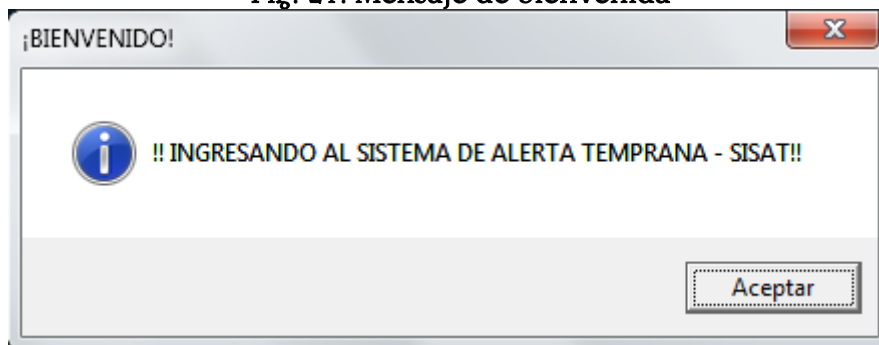
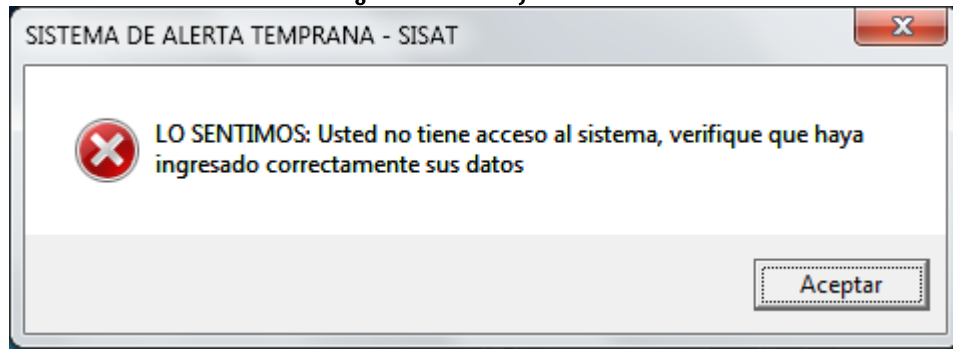




Fig. 22: Mensaje de error



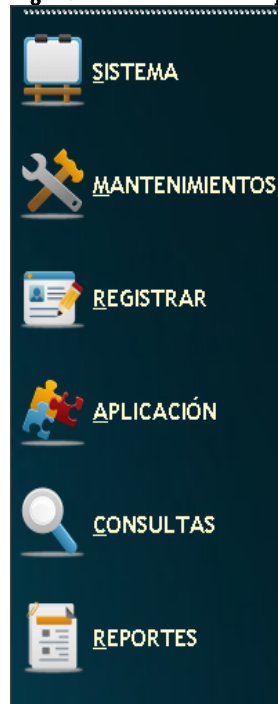
En este caso se ingresó como Administrador, entonces todas las funcionalidades del Sistema de Alerta Temprana se encontraron activas.

Podemos observar en la parte izquierda el **Menú Principal**, con todas las características del Entorno de Aprendizaje, también podemos ver que en la parte inferior se muestra el nombre de **Usuario** de la persona que ha ingresado al Sistema.

Se puede observar en la parte izquierda el **Menú Principal**, con todas las características del Sistema de Alerta Temprana, también se puede visualizar en la parte inferior se muestra el nombre de **Usuario** de la persona que ha ingresado al Sistema y el **Estado** referido al permiso que tiene en el sistema para ingresar.

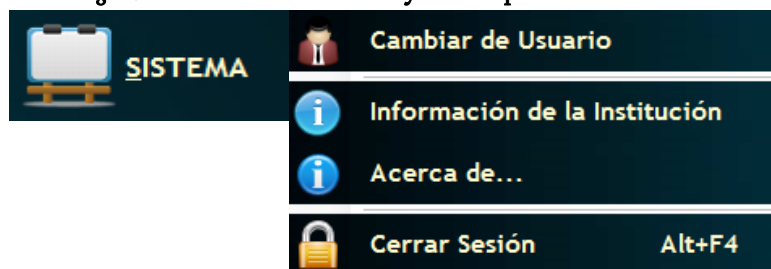
El **Menú Principal** está constituido por 6 submenús.

Fig. 23: Menú Principal



El Primer Submenú **Sistema**, consta a su vez de 4 opciones.

Fig. 24: Pestaña Sistema y sus Opciones



La opción **Cambiar de Usuario**, va permitir al usuario cerrar su sesión para ingresar con una sesión nueva sin tener que reiniciar el sistema.

La opción **Información de la Institución**, va permitir al usuario tener una pequeña información, a manera de resumen acerca de la Institución donde se implantó el Sistema de Alerta Temprana.

La opción **Acerca De**, brinda información general acerca del Sistema de Alerta Temprana como por ejemplo el lenguaje de programación con el que ha sido desarrollado, el gestor de base de datos con el que trabaja, etc.

La opción **Cerrar Sesión**, permite al usuario salir del Sistema de Alerta Temprana.

La segunda pestaña llamada **Mantenimientos**, consta a su vez de 8 opciones.

Fig. 25: Pestaña Mantenimiento y sus Opciones



La opción **Agravado**, permite al usuario registrar, modificar y buscar un agraviado, no presenta la opción eliminar pues a pesar que una denuncia haya sido resuelta siempre debe quedar la evidencia.

En un inicio tiene todas las opciones disponibles, inclusive la opción apoderado en caso de ser el **Agraviado** un menor de edad (Fig. 26), sin embargo si el **Agraviado** es mayor de edad no presenta esta opción (Fig. 27).

Fig. 26: Mantenimiento Agraviado

**Mantenimiento - AGRAVIADO**

**Información del Agraviado**

CÓDIGO: 37      OCUPACIÓN: ESTUDIANTE

NOMBRES: SANDRA      ESTADO CIVIL: SOLTERO

APELLIDOS: FLORES QUIROZ      NIVEL DE INSTRUCCIÓN: PRIMARIA COMPLETA

DNI: 47307040      ESTADO ENCONTRADO: CONCIENTE

EDAD: 17      APODERADO: ☒ SÍ   ☐ NO

GÉNERO: FEMENINO      DENUNCIANTE: ☒ SÍ   ☐ NO

Buttons: Buscar, Nuevo, Modificar, Apoderado

**Buscar Agraviado**

Acceptar

CÓDIGO	NOMBRE	APELLIDOS	EDAD	OCUPACIÓN	DNI
34	PAULA	CAPUÑAY CASTRO	39	VENDEDORA AMBULANTE	16628256
35	MARÍA	MONTENEGRO VELA	23	AMA DE CASA	44155482
36	ALEIDA	RODRIGO BAUTISTA	32	SECRETARIA	10477465
37	SANDRA	FLORES QUIROZ	17	ESTUDIANTE	47307040
38	DAVID	ALVARADO LEÓN	28	CONTADOR PÚBLICO	40909505
39	MARINA	PARDEFS GONZÁLES	51	AMA DE CASA	16545721

Fig. 27: Mantenimiento Agraviado sin registro de apoderado

**Mantenimiento - AGRAVIADO**

**Información del Agraviado**

CÓDIGO: 35      OCUPACIÓN: AMA DE CASA

NOMBRES: MARÍA      ESTADO CIVIL: SOLTERO

APELLIDOS: MONTENEGRO VELA      NIVEL DE INSTRUCCIÓN: PRIMARIA INCOMPLETA

DNI: 44155482      ESTADO ENCONTRADO: CONCIENTE

EDAD: 23      APODERADO: ☐ SÍ   ☒ NO

GÉNERO: FEMENINO      DENUNCIANTE: ☐ SÍ   ☒ NO

Buttons: Buscar, Nuevo, Modificar, Apoderado (disabled)

**Buscar Agraviado**

Acceptar

CÓDIGO	NOMBRE	APELLIDOS	EDAD	OCUPACIÓN	DNI
34	PAULA	CAPUÑAY CASTRO	39	VENDEDORA AMBULANTE	16628256
35	MARÍA	MONTENEGRO VELA	23	AMA DE CASA	44155482
36	ALEIDA	RODRIGO BAUTISTA	32	SECRETARIA	10477465
37	SANDRA	FLORES QUIROZ	17	ESTUDIANTE	47307040
38	DAVID	ALVARADO LEÓN	28	CONTADOR PÚBLICO	40909505
39	MARINA	PARDEFS GONZÁLES	51	AMA DE CASA	16545721

La opción **Imputado**, permite al usuario registrar, modificar y buscar un imputado, no presenta la opción eliminar pues a pesar que una denuncia haya sido resuelta siempre debe quedar la evidencia.

Fig. 28: Mantenimiento Imputado

**Información del Imputado**

Código: 2

Nombres: SEGUNDO VÍCTOR

Apellidos: OLIVOS TARRILLO

DNI: 28456712

Edad: 27

Género: MASCULINO

Ocupación: DESEMPLEADO

Estado Civil: CONVIVIENTE

Nivel de Instrucción: SIN INSTRUCCIÓN

Condición Anímica: ECUÁNIME

Reincidencia: REITERANTE

Descripción Física: CONTEXTURA GRUESA, DE MEDIANA ESTATURA

Buscar

Nuevo

Modificar

**Buscar Imputado**

Aceptar

CÓDIGO	NOMBRES	APELLIDOS	DESCRIPCIÓN FÍSICA	DNI	EDAD	OCUPACIÓN
1	RICARDO MIGUEL	ROMERO PERALTA	CONTEXTURA NORMAL, TEZ MORENA	27084...	36	CHOFER
2	SEGUNDO VÍCTOR	OLIVOS TARRILLO	CONTEXTURA GRUESA, DE MEDIANA ESTATURA	28456...	27	DESEMPLEADO
3	RIGOBERTO	HEREDIA PULSE	CONTEXTURA GRUESA, ALTO	15728...	69	COMERCIANTE
4	FABIO	GUERRERO VILLALOBOS	CONTEXTURA NORMAL, TEZ CLARA	35974...	30	COMERCIANTE

La opción **Apoderado**, permite al usuario registrar, modificar y buscar un apoderado, cuando la víctima es menor de edad no presenta la opción eliminar pues a pesar que una denuncia haya sido resuelta siempre debe quedar la evidencia.

Fig. 29: Mantenimiento Apoderado

**Información del Apoderado**

CÓDIGO: 1

NOMBRES: ALBINA

APELLIDOS: QUIROZ ROJAS

DNI: 80212393

OCUPACIÓN: AMA DE CASA

EDAD: 46

ESTADO CIVIL: CONVIVIENTE

NIVEL DE INSTRUCCIÓN: SECUNDARIA INCOMPLETA

PARENTESCO: MADRE

GÉNERO: FEMENINO

Buscar

Nuevo

Modificar

**Buscar Apoderado**

QUIROZ

CÓDIGO	NOMBRES	APELLIDOS	EDAD	SEXO	OCUPACIÓN	DNI
1	ALBINA	QUIROZ ROJAS	46	1	AMA DE CASA	80212393

La opción **Policía**, permite al usuario registrar, modificar y buscar un policía, no presenta la opción eliminar pues a pesar que una denuncia haya sido resuelta siempre debe quedar la evidencia.

Fig. 30: Mantenimiento Policía

**Información del Policía**

Código: 6 Estado: SUBALTERNO

Nombres: EDUARDO Comisaría: DEL NORTE

Apellidos: ESPINOZA MORE Usuario: EESPINOZA

Función: SUB-OFICIAL Contraseña: 342567

**Buscar Policía**

CÓDIGO	NOMBRES	APELLIDOS	FUNCIÓN	USUARIO	CONTRASEÑA
2	HERNÁN	BERRÍOS PÉREZ	SUB-OFICIAL TÉCNICO DE 1ra.	HBERRIOS	234567
3	JORGE	TORRES VILLEGAS	SUB-OFICIAL TÉCNICO DE 1ra.	JTORRES	345678
4	EVERT	DÍAZ BUSTAMANTE	SUB-OFICIAL DE 3ra.	EDIAZ	213456
5	SOCORRO DEL PILAR	HERNÁNDEZ SÁNCHEZ	SUB-OFICIAL TÉCNICO DE 2ra.	SHERNANDEZ	897654
6	EDUARDO	ESPINOZA MORE	SUB-OFICIAL	EESPINOZA	342567
7	JHON	SERRATO SERQUÉN	SUB-OFICIAL TÉCNICO DE 2ra.	JSERRATO	547890

La opción **Estado**, permite al usuario visualizar los estados o los niveles de accesos que existen con respecto al Sistema de Alerta Temprana, además permite crear, modificar o eliminar algún tipo de estado.

Fig. 31: Estado o Nivel de acceso sobre el Sistema de Alerta Temprana

**Información del Estado**

Código: 1 Nombre: ADMINISTRADOR

**Estado**

CÓDIGO	NOMBRE
1	ADMINISTRADOR
2	SUBALTERNO

La opción **Unidad de Intervención**, permite al usuario visualizar una Unidad de Intervención, además modificar o crear una nueva, no presenta la opción eliminar pues a pesar que una denuncia haya sido resuelta siempre debe quedar la evidencia.

Fig. 32: Unidad de Intervención

**MANTENIMIENTO - UNIDAD**

**Información de la Comisaría**

Código: 7

Comisaría: DE MUJERES Y MENORES

Ubicación: AV. SAENZ PEÑA # 495

**Comisarias**

CÓDIGO	COMISARÍA	DIRECCIÓN
3	CAMPODÓNICO	AV. JORGE CHÁVEZ # 801
4	DEL NORTE	AV. FRANCISCO CÚNEO # 1.
5	ATUSPARIA	CA. VIRREY TOLEDO # 1099

**Buttons:** + Nuevo, ↻ Modificar

**Image:** A photograph of a police station building with a car parked in front.

La opción **Delito**, permite al usuario registrar, modificar y buscar un Delito, no presenta la opción eliminar pues a pesar que una denuncia haya sido resuelta siempre debe quedar la evidencia.

Fig. 33: Delito

**MANTENIMIENTO - DELITO**

**Información del Delito**

Tipo Delito: C. EL PATRIMONIO

Código: 14

Delito: HURTO SIMPLE Y HURTO AGRAVADO

**Buttons:** Enviar, + Nuevo, ↻ Modificar

**Buscar Delito**

Search bar: [ ]

CÓDIGO	NOMBRE DEL DELITO
14	HURTO SIMPLE Y HURTO AGRAVADO
15	ROBO SIMPLE Y ROBO AGRAVADO
16	ABIGEATO
17	APROPIACIÓN ILÍCITA

**Image:** A 'No Entry' sign with a cartoon character of a person with a gun.

La opción **Tipo Delito** permite al usuario registrar, modificar y buscar un Tipo de Delito, no presenta la opción eliminar pues a pesar que una denuncia haya sido resuelta siempre debe quedar la evidencia.



Fig. 34: Tipo delito

MANTENIMIENTO - TIPO DE DELITO

**Información del Tipo Delito**

Código: 4

Tipo Delito: C. EL PATRIMONIO

**Buscar Tipo Delito**

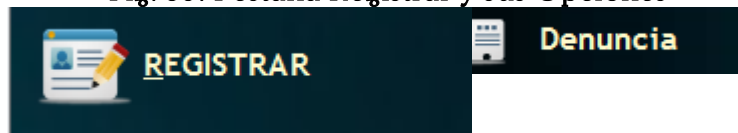
CÓDIGO	NOMBRE
1	CVCS
2	C. LA FAMILIA
3	C. LA LIBERTAD

Nuevo

Modificar

La tercera pestaña llamada **Registrar**, consta a su vez de 1 opción.

Fig. 35: Pestaña Registrar y sus Opciones



La opción **Denuncia** permite registrar una denuncia, conocer todos los involucrados con la denuncia, agraviado, imputado, la comisaría que registró lo sucedido, el (los) policías encargados de la denuncia (**Figura 36**), además cuenta con la opción **Hecho** la cual permite conocer al detalle del acto ocurrido (**Figura 37**).

La etapa de registrar denuncias consiste en la primera etapa de la construcción del Sistema de Alerta Temprana, en esta parte se lleva a cabo el proceso de sistematización de los actos delictivos (Denuncia), así como también los detalles relevantes de la persona imputada (Delincuente), lo cual posteriormente sirvió de colaboración para lograr la obtención de los perfiles delictivos.

Fig. 36: Denuncia

Fig. 37: Hecho

La cuarta pestaña y la más importante llamada **Aplicación**, consta a su vez de 2 opciones.

Fig. 38: Pestaña Aplicación y sus Opciones



La opción **Crear Dataset**, permite obtener una colección de datos, es decir filtra las características resaltantes de los implicados en un delito (variables) las que posteriormente fueron analizadas y utilizadas para definir los perfiles delictivos (Figura 39).



Los datos fueron exportados a una hoja de Excel para su análisis a través de la herramienta SPSS (Figura 40).

Fig. 39: Características Delictivas

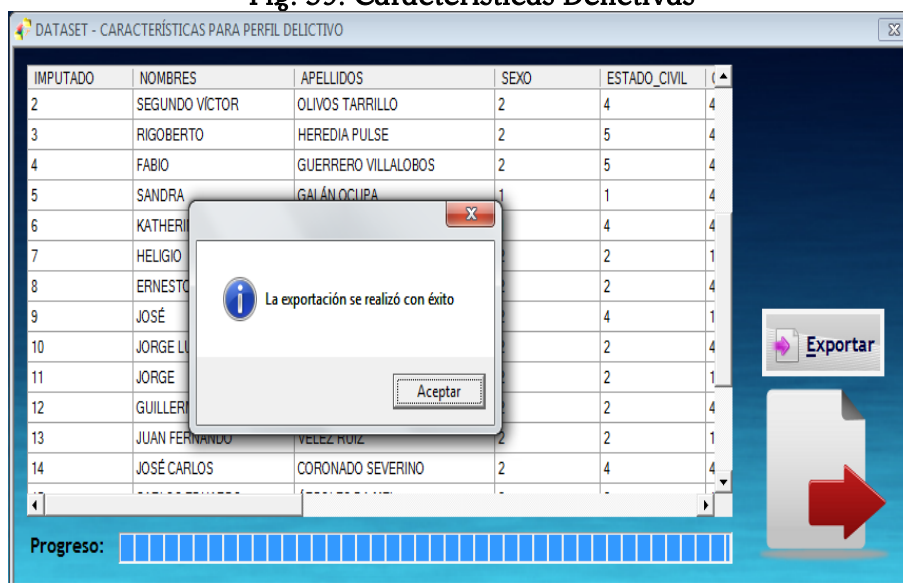


Fig. 40: Resultados de la exportación de las Características Delictivas

REGION POLICIAL DE LAMBAYEQUE							
Sistema # 000 - Chiclayo							
SISTEMA ALERTA TEMPRANA							
REPORTE DE DATASET							
Fecha de Emisión: 15/09/2012							
IMPUTADO	NOMBRES	APELLIDOS	SEXO	FECHA DEL HECHO OCURRIDO	TIPO DE AGRESIÓN	DELITO COMETIDO POR CIRCUNSTANCIA	DELITO
1	RICARDO MIGUEL	ROMERO PERALTA	MASCULINO	14/06/2009 10:30:00 P.M.	FISICA	NO	14
2	SEGUNDO VÍCTOR	OLIVOS TARRILLO	MASCULINO	14/06/2009 11:00:00 A.M.	FISICA	NO	14
3	RIGOBERTO	HEREDIA PULSE	MASCULINO	19/11/2009 02:40:00 P.M.	PSICOLOGICA	NO	33
4	FABIO	GUERRERO VILLALOBOS	MASCULINO	24/10/2009 11:00:00 P.M.	FISICA/PSICOLOGICA	NO	33
5	SANDRA	GALÁN OCUPA	FEMENINO	15/11/2009 09:40:00 P.M.	FISICA/PSICOLOGICA	NO	33
6	KATHERINE	FOELSTER SILVA	FEMENINO	14/11/2009 03:00:00 P.M.	FISICA	NO	33
7	HELIGIO	SAAVEDRA TISNADO	MASCULINO	14/12/2009 06:10:00 P.M.	FISICA/PSICOLOGICA	NO	33
8	ERNESTO	SANCHEZ LÓPEZ	MASCULINO	16/12/2009 12:35:00 P.M.	FISICA	NO	33
9	JOSÉ	CORONEL BUSTAMANTE	MASCULINO	14/12/2009 04:00:00 P.M.	PSICOLOGICA	NO	33
10	JORGE LUIS	RAMOS CALVAY	MASCULINO	10/12/2009 08:00:00 P.M.	FISICA	NO	33
11	JORGE	APONTE RIVADENEYRA	MASCULINO	11/11/2009 08:30:00 P.M.	FISICA/PSICOLOGICA	NO	33
12	GUILLERMO	SANCHEZ ZAMORA	MASCULINO	11/11/2009 12:20:00 P.M.	FISICA	NO	33
13	JUAN FERNANDO	VELEZ RUÍZ	MASCULINO	01/11/2009 05:30:00 P.M.	FISICA/PSICOLOGICA	NO	33
14	JOSÉ CARLOS	CORONADO SEVERINO	MASCULINO	24/10/2009 07:00:00 P.M.	FISICA/PSICOLOGICA	NO	33
15	CARLOS EDUARDO	ARBOLES DA MEL	MASCULINO	26/11/2009 08:00:00 P.M.	FISICA/PSICOLOGICA	NO	33
16	JUAN CARLOS	TOLEDO DELGADO	MASCULINO	24/10/2009 12:20:00 P.M.	FISICA/PSICOLOGICA	NO	33
17	VÍCTOR MARCO	VELÁSQUEZ ASTUJAMÁN	MASCULINO	14/10/2009 05:00:00 P.M.	PSICOLOGICA	NO	33

Después de obtener el dataset a través de la exportación, los datos son procesados y agrupados a través del SPSS, utilizando la técnica k-means de minería de datos, obteniéndose 12 grupos lo cual fue explicado anteriormente, en etapas iniciales del desarrollo de la Metodología.

Es ahí que empieza a funcionar el Matlab, para ello las variables que fueron definidas como entradas son obtenidas de los registros definidos en la etapa mencionada anteriormente.

Aquí se busca el implicado y se procede a calcular su grupo sospechoso al que pertenece.

La opción **Perfiles Delictivos** permite visualizar los grupos sospechosos, los datos fueron importados al sistema como parte del proceso de la Red Neuronal creada por el Matlab.

Fig. 41: Perfiles Delictivos



NOMBRES	APELLIDOS	SEXO	ESTADO CIVIL	NIVEL DE INSTRUCCIÓN	REINCIDENCIA
AURELIO	PIEDRAS PACHERRES	2	2	6	2
JUAN	FLORES FLORES	2	4	6	3
ADAN	ABANTO IDROGO	2	3	6	2
PAÚL HENRY	... SAAVEDRA CHIRA	... 2	1	4	2
ELVA	... NEYRA DE LA CRUZ	... 1	4	4	1
LEYLA	... CASTRO OBANDO	... 1	2	3	1
JUAN FERNANDO	... VÉLEZ RUIZ	... 2	2	5	1
JOSÉ	... CORONEL BUSTAMANTE	... 2	4	5	1
ROBERTO	... DAMIÁN BALAREZO	... 2	4	3	1
SEGUNDO VÍCTOR	... OLIVOS TARRILLO	... 2	4	1	1
SANTOS HERNANDEZ	... SANCHEZ GONZALEZ	... 2	2	2	2

Finalmente se muestran las personas más cercanas a un perfil delictivo según un grupo característico.

Fig. 42: Delincuente Transgresor Semiprofesional



**DELINCUENTE TRANSGRESOR SEMIPROFESIONAL**

Son delincuentes especializados en toda forma de violencia, psíquica, física o sexual. Incluye todo acto violento o de intimidación en la víctima. Este tipo de delincuentes se encuentra en cualquier nivel socioeconómico concentrándose en personas del sexo masculino, quienes también han buscado sustraer bienes o dinero. Generalmente no utilizan un arma pero si ocasionan agresiones con mayor frecuencia agresiones físicas y psicológicas ambas a la vez.

NOMBRES	APELLIDOS
AURELIO	PIEDRAS PACHERRES
JUAN	FLORES FLORES
ADAN	ABANTO IDROGO
PAÚL HENRY	... SAAVEDRA CHIRA

Fig. 43: Delincuente de Violencia Doméstica

PERFILES DELICTIVOS

DELINCUENTE DE VIOLENCIA DOMÉSTICA

Delinquentes especializados en toda forma de violencia, psíquica y/o física, perpetrada en el seno familiar. Utilizan actos de intimidación en relaciones personales de convivencia, divorcio o separación. Este tipo de agresores se centran en mayor número en personas del sexo masculino, quienes se encuentran en estado de ebriedad y/o alterados al momento de cometer este tipo de acto, no utilizan armas pero ocasionan agresiones físicas y/o psicológicas

NOMBRES	APELLIDOS
ELVA	NEYRA DE LA CRUZ
LEYLA	CASTRO OBANDO
JUAN FERNANDO	VÉLEZ RUIZ
JOSÉ	CORONEL BUSTAMANTE

Fig. 44: Ladrones Merodeadores

PERFILES DELICTIVOS

LADRONES MERODEADORES

El merodeador caracterizado porque aplica más a menudo la investigación detrás de la planificación y organización de un robo. El merodeador es como el de más edad, más maduro, que con el pasar del tiempo, decide ser más organizado a fin de maximizar sus ganancias de sus robos. Esto les obliga a estudiar su objetivo, la rutina de la víctima, sus hábitos, horarios y detalles de sus alrededores, en lo posible roban en casas sin moradores

NOMBRES	APELLIDOS
SANTOS HERMES	SANCHÉZ CORONEL
VÍCTOR RAÚL	SILVA BRAVO
JESÚS ALBERTO	MALCA SILVA
BREMILDA	TIQUILLAHUANCA ASENCIO

Fig. 45: Delincuente Profesional Consumado

PERFILES DELICTIVOS

DELINCUENTE PROFESIONAL CONSUMADO

Este tipo de delincuente se dedica a operaciones de atraco, robo y otras formas directas de rapiña en contra de las personas o propiedad ajena de forma más elaborada, por lo general actúan en grupos y tienen un espía. Tal es la habilidad de estos individuos que rara vez se ven obligados a utilizar amenazas o violencia física, actúan por lo general en estado de sobriedad. Los transgresores que forman este tipo proceden de zonas urbanas de clase baja

NOMBRES	APELLIDOS
MARCO VINICIO	PEREDA SANCHEZ
CÉSAR AUGUST...	CHUZON LAGOS ...
ABSALÓN ...	FERNÁNDEZ GUEVA...
HENRY ...	DÍAZ ORTIZ ...

Fig. 46: Semiprofesional del Delito en Contra del Bien Ajeno

PERFILES DELICTIVOS

SEMIPROFESIONAL DEL DELITO EN CONTRA DEL BIEN AJENO

Se dedican a los atracos ocasionalmente a mano armada, robos con escándalo y otras formas similares de delitos en contra de la propiedad y/o las personas, la estrategia que se sirven es relativamente simple y poco elaborada. El criminal suele actuar en forma directa y burda, por ello en algunos casos incurre en agresiones a la víctima

NOMBRES	APELLIDOS
JHON IVAN	RAMOS SALAZAR
WILSON	LOZANO VASQUEZ
ELIZABETH	CALDERÓN DAGA
VÍCTOR HUGO	OYOLA VILCHEZ

Fig. 47: Transgresores Familiares

PERFILES DELICTIVOS	
TRANSGRESORES FAMILIARES	
<p>Delinquentes que atentan sobre el bienestar familiar, con mayor frecuencia bajo estado de sobriedad aunque algunos en estado de ebriedad. Los transgresores familiares generalmente no hacen uso de un arma, sin embargo cuando utilizan un arma es común que sea arma blanca, además principalmente buscan disminuir a la persona (Agresión psicológica)</p>	
NOMBRES	APELLIDOS
MARÍA BRIGIDA ...	PALOMINO IZQUIERD...
BETTY ...	VIDAURE FARRO ...
ANDY WILSON ...	CORONEL CRUZADO...
LUIS ALBERTO ...	PANDURO FERRÉ ...

Fig. 48: Lesionador Indolente

PERFILES DELICTIVOS	
LESIONADOR INDOLENTE	
<p>Delinquentes caracterizados por un bajo control de sus impulsos, pudiendo alcanzar rápidamente niveles altos de agresividad y violencia, son generalmente varones alterados o bajo efectos de sustancias nocivas, no se conmueven ante el dolor ajeno por ello son altamente reincidentes, además la posibilidad de rehabilitación es baja debido a que sus procesos resilientes se encontrarían en un nivel de desarrollo mínimo</p>	
NOMBRES	APELLIDOS
ADRIEL ENRIQUE	BERNAL PEREZ
JHONATHAN	BOCANEGRA CARRA...
WILDER	MUÑOZ PAISIG
MOISÉS	SILVA FLORES

Fig. 49: Ofensor por Impulsividad

PERFILES DELICTIVOS

OFENSOR POR IMPULSIVIDAD

Delincuente ofensor, aprovecha la oportunidad que se le presente, comete su delito sin ser planificado, actúa impulsado por arrebatos de celos, odio o venganza. Generalmente su agresión se ve detonada por el consumo de drogas u alcohol, o bien, por un trastorno mental agudo o transitorio. Además no se ven a sí mismos como delincuentes porque no se involucran en bandas o pandillas, justifican sus actos con la idea de que tienen problemas

NOMBRES	APELLIDOS
DIEGO WILSON	ZURITA CHOPITEA
VÍCTOR ORLANDO	REYES VINCES
CRONWEL ARTI...	UGAZ PAZ
EUSEBIO	ALCALDE LOZANO

Fig. 50: Agresor de Peligrosidad Extrema

PERFILES DELICTIVOS

AGRESOR DE PELIGROSIDAD EXTREMA

En su mayoría estos delincuentes se contentan con maltratar de manera física o psicológica a sus víctimas sin despojar de pertenencias ni incurrir en delitos en contra de la propiedad ajena. Principalmente poseen un nivel de educación medio-bajo. No se conceptúan a sí mismos como delincuentes, se ven como víctimas de un medio ambiente hostil y se muestran retadores

NOMBRES	APELLIDOS
JULIO CÉSAR ...	RIVERA CASTILLO
LUIS LEONIDAS	LEÓN LECA
BETTY DORIS	CHINCHAY FARRO
RICARDO REMIGIO	SEGURA FERNÁNDEZ

Fig.51: Lesionador de Personas

PERFILES DELICTIVOS

LESIONADOR DE PERSONAS

Delincuentes que cometen delitos generalmente violentos contra la integridad física de las personas. Estos ofensores accidentales no tienen formada de sí mismos una imagen de criminales y proceden de distintos sectores socioeconómicos. Además por lo general están bajo efectos de alguna sustancia nociva para la salud (alcohol, drogas, etc.)

NOMBRES	APELLIDOS
OSCAR	CARDENA PERLECHE
PILAR VERÓNICA	CARLOS TOMAY
TOMÁS	FERNÁNDEZ MONT...
SANTOS	TIQUILLAHUANCA ...

Fig. 52: Delincuente Profesional Organizado Local

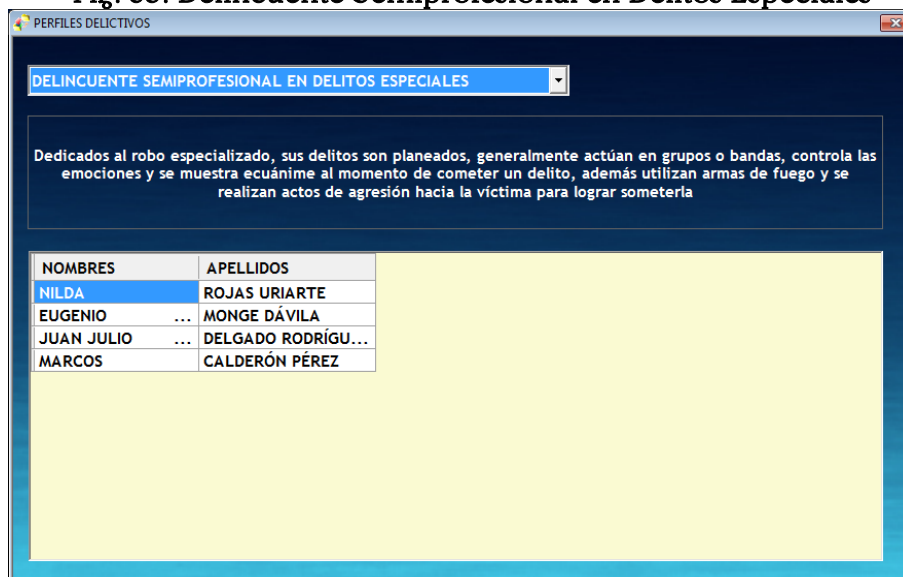
PERFILES DELICTIVOS

DELINCUENTE PROFESIONAL ORGANIZADO LOCAL

Dedicados al robo especializado, sus delitos son planeados, generalmente actúan en grupos o bandas, controla las emociones y se muestra ecuaníme al momento de cometer un delito, además utilizan armas de fuego y se realizan actos de agresión hacia la víctima para lograr someterla

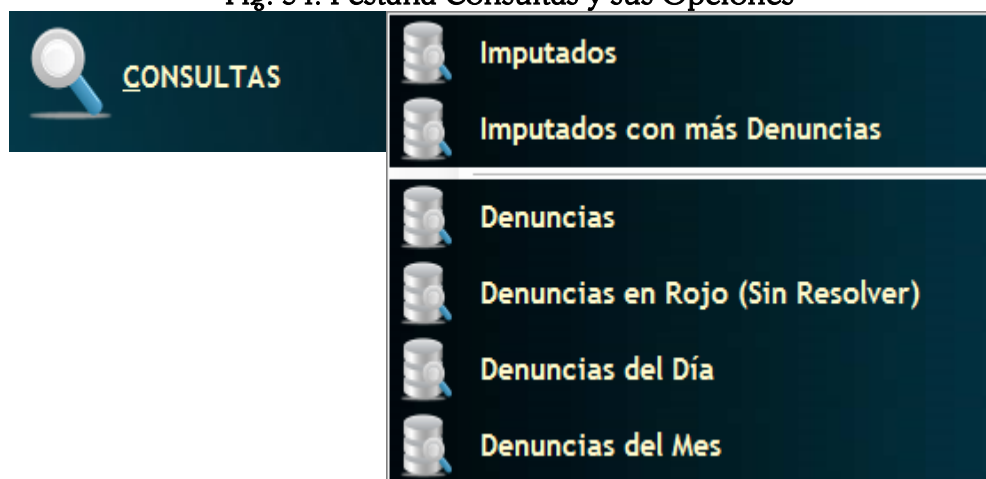
NOMBRES	APELLIDOS
MIGUEL ÁNGEL ...	FLORES JIMÉNEZ ...
JESSENIA RUBÍ ...	VINCES HINOJOSA ...
JOSÉ GUILLERMO...	HINOJOSA UCHOFE...
EVELIN MARGARE...	VINCES HINOJOSA ...

Fig. 53: Delincuente Semiprofesional en Delitos Especiales



La quinta pestaña llamada **Consultas**, consta a su vez de 6 opciones.

Fig. 54: Pestaña Consultas y sus Opciones



La opción **Consultar Imputados** permite visualizar un listado de todos los Imputados o implicados en un delito, así como también buscarlos por apellidos.



Fig. 55: Consultar Imputados

**CONSULTAR - IMPUTADOS**

Información para Consulta

☒ Por Imputado

[Ver Detalle](#)

**RESULTADO DE LA CONSULTA**

CÓDIGO	NOMBRES	APELLIDOS
1	RICARDO MIGUEL	ROMERO PERALTA
2	SEGUNDO VÍCTOR	OLIVOS TARRILLO
3	RIGOBERTO	HEREDIA PULSE
4	FABIO	GUERRERO VILLALOBOS
5	SANDRA	GALÁN OCUPA
6	KATHERINE	FOELISTER SILVA
7	HELIGIO	SAAVEDRA TISNADO
8	ERNESTO	SÁNCHEZ LÓPEZ
9	JOSÉ	CORONEL BUSTAMANTE
10	JORGE LUIS	RAMOS CALVAY
11	JORGE	APONTE RIVADENEYRA
12	GUILLERMO	SANCHEZ ZAMORA
13	JUAN FERNANDO	VELEZ RUÍZ
14	JOSÉ CARLOS	CORONADO SEVERINO

La opción **Consultar Imputados con más denuncias** permite visualizar un listado de todos los Imputados o implicados que han sido más denunciados, además permite exportar a una hoja Excel.

Fig. 56: Consultar Imputados

**CONSULTAR - DELINCUENTES CON MÁS DENUNCIAS**

Desde: jueves , 11 de febrero de 2010

hasta: martes , 11 de junio de 2013

[Consultar](#) [Exportar](#)

CÓDIGO	NOMBRE	APELLIDOS	TOTAL_DENUNCIAS
1	RICARDO MIGUEL	ROMERO PERALTA	4
2	SEGUNDO VÍCTOR	OLIVOS TARRILLO	2

La opción **Consultar Denuncias** permite visualizar un listado de todas las denuncias por estado o por el lugar donde fueron registradas (comisaría).

Fig. 57: Consultar Denuncias

**Información de la Consulta**

☒ CONSULTAR DENUNCIAS POR ESTADO: PENDIENTE

☐ CONSULTAR DENUNCIAS POR COMISARÍAS:

**RESULTADO DE LA CONSULTA**

CÓDIGO	DESTINO DE LA DENUNCIA
1	PENDIENTE DE INVESTIGACIÓN
3	PENDIENTE DE INVESTIGACIÓN
6	PENDIENTE DE INVESTIGACIÓN
7	PODER JUDICIAL
12	PENDIENTE DE INVESTIGACIÓN
14	PENDIENTE DE INVESTIGACIÓN

La opción **Consultar Denuncias en Rojo** permite visualizar un listado de todas las denuncias que están pendientes o que todavía no se han resuelto, además permite la exportación a Excel.

Fig. 58: Consultar Denuncias en Rojo

**CONSULTAR - DENUNCIAS EN ROJO**

CÓDIGO	ESTADO	FECHA	COMISARIA
81	PENDIENTE	09/03/2009 10:00 a.m.	JOSÉ LEONARDO ORTIZ
3	PENDIENTE	10/04/2009 02:00 a.m.	DEL NORTE
1	PENDIENTE	12/05/2009 04:00 p.m.	DEL NORTE
12	PENDIENTE	26/09/2009 11:00 a.m.	DEL NORTE
18	PENDIENTE	02/10/2009 02:00 p.m.	DEL NORTE
80	PENDIENTE	30/10/2009 11:00 a.m.	CAMPODÓNICO
14	PENDIENTE	30/10/2009 12:30 p.m.	DEL NORTE

**Exportar**

**Generación de Proceso**

Progreso:

Número de Registros: **10**

**Denuncias sin Resolver**

La opción **Consultar Denuncias del Día** permite visualizar un listado de todas las denuncias de acuerdo a un día específico, además permite la exportación a Excel.

Fig. 59: Consultar Denuncias del Día

**CONSULTAR - DENUNCIAS DEL DIA**

**Opciones**  
 Fecha: 12/03/2009 [Consultar] [Exportar]

**Generación del Proceso**  
 Progreso: [ ]

**RESULTADOS:**

CÓDIGO	ESTADO	DESTINO	FECHA
115	RESUELTO	JUZGADO DE FAMILIA	12/03/2009

**IMPUTADO**

NOMBRE	APELLIDOS
MIRIAM	LEON TENORIO

**AGRAVIADO**

NOMBRE	APELLIDOS
PAOLA	DELGADO AVANTO

**Denuncias del Día:**

La opción **Consultar Denuncias del Mes** permite visualizar un listado de todas las denuncias de acuerdo a un mes especificado, además permite la exportación a Excel.

Fig. 60: Consultar Denuncias del Mes

**CONSULTAR - DENUNCIAS DEL MES**

**Opciones**  
 Fecha: 11/08/2009 [Consultar] [Exportar]

**Generación del Proceso**  
 Progreso: [ ]

**RESULTADOS:**

CÓDIGO	ESTADO	DESTINO	FECHA
10	RESUELTO	PODER JUDICIAL	25/08/2009

**IMPUTADO**

NOMBRE	APELLIDOS
JORGE LUIS	RAMOS CALVAY

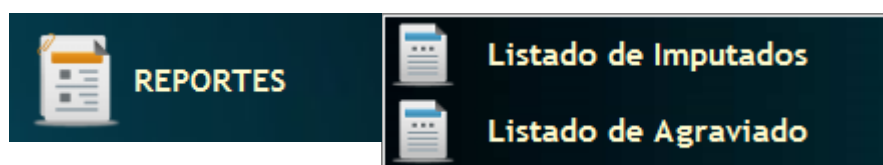
**AGRAVIADO**

NOMBRE	APELLIDOS
ORFELINDA	IPANAQUE MEDINA

**Denuncias del Mes:**

La sexta pestaña llamada **Reportes**, consta a su vez de 2 opciones.

Fig. 61: Pestaña Reportes y sus Opciones



La opción **Listado de Implicados** permite ver un reporte de todas las personas implicadas en un delito.

Fig. 62: Listado de Imputados

Fig. 02: Estado de Imputados

Sistema de Alerta Temprana

 **Región Policial Lambayeque**

---

## REPORTE - LISTA DE IMPUTADOS

INFORMACIÓN DEL IMPUTADO				
<u>CÓDIGO:</u>	<u>NOMBRES:</u>	<u>APELLIDOS:</u>	<u>EDAD:</u>	<u>OCUPACIÓN:</u>
1	RICARDO MIGUEL	ROMERO PERALTA	36	CHOFER
2	SEGUNDO VÍCTOR	OLIVOSTARRILLO	27	DESEMPLEADO
3	RIGOBERTO	HEREDIA PULSE	69	COMERCIANTE
4	FABIO	GUERRERO VILLALOBOS	30	COMERCIANTE
5	SANDRA	GALÁN OCUPA	20	ESTUDIANTE

La opción **Listado de Agraviados** permite ver un reporte de todas las personas que han sido víctimas.

Fig. 63: Listado de Agraviados

Fig. 05: Estado de Agriavados

Sistema de Alerta  
Temprana

Región Policial  
Lambayeque

## REPORTE - LISTA DE AGRAVIADOS

### INFORMACIÓN DEL AGRAVIADO

<u>CÓDIGO:</u>	<u>NOMBRES:</u>	<u>APELLIDOS:</u>	<u>DNI:</u>	<u>EDAD:</u>	<u>OCUPACIÓN:</u>
1	MARGARITA ELENA	ESPINO PARDO	27172345	31	AMA DE CASA
2	MARÍA AUDILA	VÁSQUEZ ROJAS	27437935	31	AMA DE CASA
3	DORIS YLIDILIA	ARIZAGA CHUMACERO	16457412	65	COMERCIANTE
4	OFELIA	ESPINOZA MALUQUIS	41051871	28	AMA DE CASA
5	ANA CINTHYA	PURIHUAMAN MORA	45657324	20	COMERCIANTE

#### 4.6. Resultado del Post – Test

El Post – Test constó en la toma de tiempos para controlar la demora en los procesos principales como del proceso de Registro de denuncia; se obtuvieron los siguientes resultados.

Tabla 49: Resultados obtenidos en el Post - Test

RESULTADO POST – TEST		
T01	5	Minutos
T02	4	Minutos
T03	4	Minutos
T04	6	Minutos
T05	5	Minutos
T06	7	minutos
T07	7	minutos
T08	6	minutos
T09	4	minutos
T10	7	minutos

El promedio de registro de una denuncia por parte de los policías es de 5.4 minutos.

## VII. DISCUSIÓN

En este capítulo se hizo un análisis de los resultados del Sistema de Alerta Temprana desarrollado para la Región Policial. El análisis estuvo enfocado a evaluar los indicadores planteados en el capítulo III. Los indicadores tomados en cuenta fueron:

- ❖ Índice de seguridad de acceso al sistema
- ❖ Índice de satisfacción por el uso del Sistema de Alerta Temprana
- ❖ Usabilidad del Sistema de Alerta Temprana
- ❖ Tiempo de Respuesta por consulta
- ❖ Número de reportes por personas denunciadas
- ❖ Tiempo de dedicación en el registro de la denuncia
- ❖ Tiempo en búsquedas de información por criterios
- ❖ Tiempo de demora en la emisión resolución de las denuncias
- ❖ Índice de errores en coincidencias de datos del implicado por expediente
- ❖ Costo por materiales de escritorio

Además en el presente capítulo se discuten antecedentes y los principales hallazgos de la investigación realizada.

Ratifico lo mencionado en la Investigación de Caverio (2009) que expresa que la Minería de Datos ayuda a extraer conocimiento a partir de datos que las empresas han generado producto de su negocio. Este conocimiento puede generar aplicaciones de alto valor agregado, gracias a la utilización de diversas técnicas como clustering y al entendimiento apropiado del proceso de Minería de Datos.

Afirmo lo mencionado en la Investigación de Gutiérrez (2008) que expresa que el análisis de la información criminal – delictiva es un punto de vital importancia a los fines de prevención del delito. Sin embargo es necesario hacer hincapié en la persona que ejecuta dicho delito, el delincuente, a fines de obtener patrones de comportamiento y conducta de los delincuentes, basándose en la información delictiva.

Se coincidieron en los resultados de la Investigación de Barrantes y Córdova (2007) que expresan que a través de la utilización de técnicas de Minería de Datos se pudo identificar características no observables que se producían durante la gestión de Matrículas logrando mejorar la toma de decisiones analíticas referentes a la Gestión de Matrículas de la Universidad Católica Santo Toribio de Mogrovejo. Además se logró la reducción del tiempo de acceso y repuesta en la obtención de la búsqueda de información para generar los reportes de gestión de Matrículas, en la presente tesis desarrollada también se logró la reducción del tiempo de búsquedas de información en cuanto a la obtención de patrones de comportamiento de los imputados (perfiles delictivos).

Para la evaluación, se analizaron los indicadores antes mencionados, para el registro de las denuncias y determinación de perfiles delictivos de forma manual, y para el registro de las denuncias y determinación de perfiles delictivos utilizando el sistema.

❖ **Índice de seguridad de acceso al sistema realizada a los Policías.**

**Tabla 50: Resultados de Encuesta Realizada a los Policías (Ver Anexo 05)**

ITEM	DESCRIPCIÓN	SI		NO	
		Nº	%	Nº	%
1.	La aplicación cuenta con niveles de acceso	10	100%	0	0%
2.	Existen mensajes que emite el sistema en caso de ingresar con un usuario y/o contraseña inválida	10	100%	0	0%
3.	Se le tienen prohibidas algunas opciones a la aplicación debido al nivel de acceso destinado para usted	10	100%	0	0%
4.	La aplicación se ejecuta lentamente.	3	30%	7	60%
5.	La aplicación se ha detenido inesperadamente en algún momento	1	10%	9	90%

❖ **Índice de satisfacción por el uso del Sistema de Alerta Temprana**

**Tabla 51: Resultados de Encuesta Realizada a los Policías (Ver Anexo 05)**

ITEM	DESCRIPCIÓN	SI		NO	
		Nº	%	Nº	%
1.	Disfruto del manejo del sistema	2	20%	8	80%
2.	Recomendaría su utilización dentro de la Institución donde labora	2	20%	8	80%
3.	Le gustaría usarlo diariamente	3	30%	7	70%
4	Aprender a usarlo me tomó mucho tiempo.	2	20%	8	80%
5.	Nunca tengo la información	3	30%	7	70%

	suficientemente necesaria en pantalla				
6.	Tiene una interfaz muy amigable	2	20%	8	80%

#### ❖ Usabilidad del Sistema de Alerta Temprana

**Tabla 52: Resultados de Encuesta Realizada a los Policías (Ver Anexo 05)**

ITEM	DESCRIPCIÓN	SI		NO	
		Nº	%	Nº	%
1.	Es complicado de usar	2	20%	8	80%
2.	A veces en algunos puntos, no sé cómo continuar	3	30%	7	70%
3.	Si el software termina es difícil reiniciarlo.	3	30%	7	70%
4	Solo me siento seguro usando comandos y operaciones que me son familiares.	2	20%	8	80%
5.	La organización de los menús o la información listada es poco lógica	3	30%	7	70%
6.	Aprender a usar las nuevas funcionalidades es difícil	4	40%	6	60%
7.	Los mensajes para prevenir errores no son adecuados.	2	20%	8	80%
8.	La aplicación se ejecuta lentamente	5	50%	5	50%

#### ❖ Tiempo de Respuesta por consulta

##### *Manual*

Tiempo Listado de agraviados: 5 minutos

Tiempo Listado de imputados por día: 5 minutos

Tiempo Listado de imputados por mes: 15 minutos

Tiempo en búsqueda de un caso específico: Dependiendo el tipo de caso puede demorar entre 15 minutos a 40 minutos.

##### *Sistema*

Tiempo Listado de agraviados: 1 minuto

Tiempo Listado de imputados por día: 35 segundos

Tiempo Listado de imputados por mes: 40 segundos

Tiempo en búsqueda de un caso específico: 2 minutos



### ❖ **Número de reportes por personas denunciadas**

#### *Manual*

Se genera un reporte o un registro separado por cada denuncia, depende de la cantidad de delitos que comete un implicado, se genera más de 1 reporte.

#### *Sistema*

En 1 único reporte muestra todo los detalles de la persona implicada en uno más denuncias.

### ❖ **Tiempo de dedicación en el registro de la denuncia**

#### *Manual*

Tiempo de registro de denuncia: Tiempo promedio de acuerdo a la toma de tiempo es 10.5 minutos

#### *Sistema*

Tiempo de registro de denuncia: Tiempo promedio de acuerdo a la toma de tiempo es 5.4 minutos

### ❖ **Tiempo en búsquedas de información por criterios**

#### *Manual*

Tiempo de búsqueda: 2 horas.

#### *Sistema*

Tiempo de búsqueda: 4 minutos.

### ❖ **Tiempo de demora en la emisión resolución de las denuncias**

#### *Manual*

Tiempo de demora en la emisión de resolución de denuncia en caso de falta leve: 1 a 2 días.

Tiempo de demora en la emisión de resolución de denuncia en caso de falta grave: 15 días.

#### *Sistema*

Tiempo de demora en la emisión de resolución de denuncia en caso de falta leve: 1 a 2 días.

Tiempo de demora en la emisión de resolución de denuncia en caso de falta grave: 15 días.

### ❖ **Índice de errores en coincidencias de datos del implicado por expediente**

#### *Manual*

Porcentaje de error: 2%, equivalente a 108 casos para el año 2009.

#### *Sistema*

Porcentaje de error: 0%

#### ❖ Costo por materiales de escritorio

##### *Manual*

Costo en Papel Bond anual: 171 soles

Costo de lapiceros anual: 50 soles

Costo de correctores anual: 60 soles

Costo de fólder anual: 35 soles

Costo de archivadores anual: 45 soles

Libro de Actas anual: 65 soles

**Total: 426 soles.**

#### *Sistema*

Costo de licencia Microsoft Visual Studio 2008: 280.00

Costo de Microsoft SQL Server 2005 Developer Edition:  
133.00

**Total: 413 soles.**

De acuerdo a los resultados obtenidos sobre la toma de tiempo para controlar la demora en los procesos principales como del proceso de Registro de denuncia (Tabla 53).

**Tabla 53: Tiempos registrados antes y después de la implementación del Sistema**

RESULTADO PRE - TEST			RESULTADO POST - TEST		
T01	11	Minutos	T01	5	minutos
T02	10	Minutos	T02	4	minutos
T03	9	Minutos	T03	4	minutos
T04	10	Minutos	T04	6	minutos
T05	10	Minutos	T05	5	minutos
T06	12	Minutos	T06	7	minutos
T07	12	Minutos	T07	7	minutos
T08	11	Minutos	T08	6	minutos
T09	9	Minutos	T09	4	minutos
T10	11	Minutos	T10	7	minutos

**Tabla 54: Diferencia de las Medias**

DIFERENCIA DE MEDIAS
HIPÓTESIS PROPUESTA <HIPÓTESIS ACTUAL

5.5	<	10.4
-----	---	------

En los resultados expuestos anteriormente se puede apreciar que los tiempos de demora se han reducido significativamente (4 1/5 minutos aprox.) Ello confirma la hipótesis: El desarrollo de un modelo de minería de datos como herramienta de apoyo puede contribuir a la caracterización de posibles perfiles delictivos de la ciudad de Chiclayo.

La implementación del Sistema de Alerta Temprana es viable, debido a que los costos para el desarrollo de la misma son accesibles por la empresa, además es factible, puesto que es una nueva forma de ofrecer información, la cual permite identificar perfiles delictivos de acuerdo a las características que estos poseen, con lo cual se podrá realiza proyecciones y plantear medidas de prevención.

## VI. PROPUESTA

Entre las ventajas que ofrece un Sistema de Alerta Temprana destacan: Registro de denuncias actualizadas, inexistencia de duplicidad de datos, Búsquedas de historiales delictivos más rápidos, pues el registro que se lleva actualmente en las Comisarías de la Ciudad de Chiclayo se realiza de manera tradicional, es decir registro en el libro de actas, o registro a través de máquina de escribir, en el cual cada nuevo caso es registrado sin tener en cuenta que un mismo implicado puede haber sido registrado con anterioridad, además de esto no se incurre en gastos de materiales como papel, libros, cuartos donde se guardan todas las actas de las denuncias.

Con respecto a la Infraestructura Física, de la Región Policial cuenta con 3 computadoras en el área de estadística, referido al registro de denuncias, pero se recomienda cambiarlas para cubrir los requerimientos del sistema, además una de ellas se encuentra obsoleta con lo cual obstruyen el correcto funcionamiento, de manera que se genere el ambiente adecuado para que cada usuario pueda interactuar de manera eficaz, sin interrupción con el Sistema de Alerta Temprana

En el área de Estadística las debe contar con mínimo con 5 computadoras, que presenten las siguientes características:

1. Plataforma: x32 – x64.
2. Memoria RAM: 512 MB a más.
3. Procesador: Pentium IV a más.
4. Disco Duro: las computadoras deben tener como mínimo 4GB de espacio disponible.

Además se necesita que las computadoras tengan como Sistema Operativo, Windows ya sea XP o cualquier versión posterior; y por último que tengan instalado Microsoft SQL Server Developer 2005, ya que este programa permitirá el manejo y administración de la base de datos y el almacenamiento de los datos que se registren a través de la aplicación.

Así mismo se necesita la instalación del Matlab2010A en las computadoras, ya que este programa permite la clasificación de los perfiles delictivos a través de la red neuronal creada.

Debido a la complejidad de la aplicación se requiere la inversión de recursos financieros empezando por los costos de implementación de mejoras a la infraestructura tecnológica presente, costos de implementación y mantenimiento de la aplicación, costos de configuración y administración. También se requiere invertir en recursos humanos, es decir ofrecer capacitación a los policías sobre esta nueva herramienta, para que se familiaricen y logren utilizarla con la mayor una facilidad posible. Además se requiere de personal técnico que asegure el buen funcionamiento de la aplicación.

Se propone que el Sistema de Alerta Temprana sea utilizado por los policías como una herramienta de apoyo en el Proceso de Registro de denuncias y caracterización de Perfiles Delictivos, que permita a los policías reducir el foco de búsqueda de los delincuentes que han cometido un delito.

La información contenida por el Sistema de Alerta Temprana se tomó en cuenta de la Documentación de la Región Policial Lambayeque y actas de denuncias, la cual contiene la información relevante de un hecho ocurrido así como también la información del implicado, con la finalidad de detectar perfiles delictivos se propone que el Sistema de Alerta Temprana sea aplicado dentro de las comisarías, ya que como se pudo comprobar esta herramienta apoya a reducir tiempos de pérdida en registro y actualización de denuncias, además a realizar búsquedas de perfiles en tiempos reducidos.

Si bien el Sistema de Alerta Temprana cumple con las características propuestas, se puede aumentar estas mismas y lograr una mayor funcionalidad y nuevos usos como es el extender la funcionalidad a un número mayor de delitos puesto que el sistema solo abarca los Delitos contra el Patrimonio, contra la vida el cuerpo y la salud y contra la familia.

## VII.CONCLUSIONES

Mediante la aplicación del software desarrollado en la presente tesis en la Región Policial Lambayeque, se concluye que:

1. Se logró la identificación de patrones delictivos con el apoyo del Sistema de Alerta Temprana basado en la Minería de Datos desarrollado para la Región Policial lográndose definir 12 grupos delictivos, con distintas características y conductas permitiendo validar conocimientos preexistentes. Además se logró caracterizar a los implicados en un delito en base a sus atributos más relevantes.
2. La implementación del Sistema de Alerta Temprana demostró su eficiencia pues se contrastó con la información obtenida anteriormente a la implementación en la Región Policial, dando como resultado que el sistema redujo considerablemente el tiempo dedicado en 4.9 minutos en el proceso de registro de denuncias y en más de 1 hora y media en el proceso de búsquedas de perfiles delictivos.
3. El sistema permitió mantener la información ordenada y actualizada, de tal manera que pueda acceder a la información de manera rápida, logrando reducir en promedio 11 minutos en cuanto a las búsquedas manuales.
4. Los costos de materiales destinados al registro de denuncias se redujeron de manera poco notoria, pues a pesar de ser una herramienta informática se incurrió en gastos de licencias lo cual solo marcó una diferencia mínima de 13 soles con respecto a los gastos de materiales.
5. Finalmente, se alcanzó satisfacer las necesidades de los policías en un 77%, quienes inicialmente presentaban oposición al cambio, pues se demostró que esta nueva herramienta sirve como apoyo en el proceso tradicional de registro de denuncias y caracterización de perfiles delictivos reduciendo el foco de búsqueda de implicados en un hecho.

## VIII. REFERENCIAS BIBLIOGRÁFICAS

- Acosta. s.f. *Minería de Datos y Descubrimiento de Conocimiento (Data Mining and Knowledge Discovery)*. La Habana: Facultad de Economía, Universidad de La Habana.
- Barrantes y Córdova. 2007. *Aplicación de una Solución Business Intelligence con Minería de datos para la mejora de toma de decisiones en la Gestión de Matrículas de la Universidad Católica Santo Toribio de Mogrovejo*, Lambayeque – Perú. Chiclayo: USAT.
- Britos. et.al. 2008. *Explotación de Información Aplicada a Inteligencia Criminal en Argentina*. Argentina: Centro de Ingeniería del Software e Ingeniería del Conocimiento. Universidad de Buenos Aires.
- Caverro y Sepúlveda. 2009. *Aplicación de Minería de Datos para la Detección de Anomalías: Un Caso de Estudio*. Chile: *WORKSHOP INTERNACIONAL EIG2009*. Universidad de La Frontera. [http://sistemas.dis.ufro.cl/eig2009/Documentos/Art\\_8.pdf](http://sistemas.dis.ufro.cl/eig2009/Documentos/Art_8.pdf). (Accedido Junio 15, 2010)
- Chanduví. et.al. 2004. *Estudio de opinión: Seguridad ciudadana en la Provincia de Chiclayo*. Centro de Opinión Pública (COP – USS). Chiclayo - Perú: Universidad Señor de Sipán.
- Chen, H., Z. Daniel, A. Homa, W. Wojciech, y S. Jenny. 2003. *COPLINK Managing Law Enforcement Data And Knowledge*. USA: The University of Arizona.
- Criado. 2010. «La policía del 'precrimen' ya está en marcha: Un software predice si un joven delincuente puede reincidir.» *Diario Público*, 9 de mayo. Madrid. España. <http://www.publico.es/311261>. (Accedido Mayo 25, 2010)
- Daza. 2007. *Métodos para mejorar la calidad de un conjunto de datos para descubrir conocimiento*. Puerto Rico: Ciencias e Ingeniería de la Información y Computación.
- Díaz 2007. *El Nuevo paradigma de la Investigación Criminal*. México. Chihuahua: Gobierno del Estado.
- Domínguez s.f. *La Policía de Investigación Criminal*. Venezuela: Memorias.
- Gómez y Suárez. 2004. *Sistemas de información: herramientas prácticas para la gestión empresarial*. México, D.F.: Alfaomega.

- Gutiérrez. 2008. *Caracterización de la Población Carcelaria en Argentina mediante la aplicación de Minería de Datos para la prevención de hechos delictivos*. Argentina: Instituto Tecnológico de Buenos Aires - ITBA.
- Gutierrez.2000. *Introducción a la Inteligencia Artificial. Aplicaciones*. España: Dpto. de Matemática Aplicada y Ciencias de la Computación. Universidad de Cantabria
- Hernández, Ramírez y Ferri. 2004. *Introducción a la Minería de Datos*. Madrid: PEARSON EDUCACIÓN, S.A.
- Hernández, Ramírez, y Ferri 2004 citando a:
  - Deerwester et al. 1990. *Indexing by latent semantic analysis, journal of the American Society*. For information Science.
  - Etzioni. 1996. *The World Wide Web: quagmire or gold mine? Communications of the ACM*.
  - Lagus et al. 1999. *WebSom for textual data mining. Artificial Intelligence Review*, vol. 13, Kluwer Academic Publisher.
  - Sahami et al. 1996. *Applying the Multiple Cause Mixture Model to Text Categorization, in Proceeding of AAAI Spring Symposium on machine Learning and Information Access*.
  - Witten & Frank. 2000. *Data Mining. Practical Machine Learning Tools and Techniques with java Implementations*. Morgan Kaufmann Publishers.
- Hilera y Martínez 2000. *Redes Neuronales Artificiales*. Bogotá, D.C. : Alfaomega,
- II Dirección Territorial Chiclayo. 2008 «Plan Operativo 'Retén 2008'. Delincuencia.» Chiclayo. Perú. [http://www.pnp.gob.pe/logros/direcc\\_territ/2\\_chiclayo.htm](http://www.pnp.gob.pe/logros/direcc_territ/2_chiclayo.htm). (Accedido Junio 01, 2012)
- Instituto Nacional de Estadística e Informática (INEI). 2013. “Perú: Perfil del Trabajo Infantil al 2013”.Perú.
- Instituto de Formación de Adolescentes y Niños Trabajadores - INFANT. 2015. Citando a INEI 2011. En el Perú hay más niños trabajando que en cualquier otro país de Sudamérica. Perú. <http://www.infant.org.pe/en-el-peru-hay-mas-ninos-trabajando-que-en-cualquier-otro-pais-de-sudamerica/>(Accedido Junio 15, 2015)
- Martínez. 2009. *Caracterización y predicción espacio-temporal de patrones delictivos mediante modelos lógicos-combinatorios*. México, D.F.: Instituto Politécnico Nacional.
- Martínez. 2009 citando a:



- Montiel. 1997. *Criminalística*. México: Limusa Noriega Editores.
- Osborne y Wernicke. 2006. *Introduction to Criminal Analysis*. Estados Unidos: Routledge.
- Mayhuay. 2006. *Identificación y Predicción de Intereses de los alumnos de una facultad de ingeniería aplicando técnicas de agrupamiento*. Lima: Universidad Nacional de Ingeniería.
- Müller 2006. *La policía Nacional del Perú y la investigación criminal en el proceso penal acusatorio*. Perú: Universidad Cesar Vallejo de Trujillo.
- Oyarzo. S.f. *Introducción a la Informática*. Madrid: Departamento de Automática y Ciencias de la Computación. Universidad de Alcalá.
- Palma y Marín 2008. *Inteligencia Artificial: Métodos, técnicas y aplicaciones*. Madrid: Mc Graw-Hill Interamericana.
- Pérez 2007. *Minería de Datos: Técnicas y herramientas*. México D.F: Thomson.
- Pérez y Santín. 2006. *Data mining: soluciones con Enterprise Miner*. México, D.F.: Alfaomega.
- Perversi. 2007. *Aplicación de Minería de Datos para la exploración y detección de patrones delictivos en Argentina*. Buenos Aires: Instituto Tecnológico de Buenos Aires.
- Vallejos. 2007. *Minería de Datos*. Argentina: Facultad de Ciencias Exactas, Naturales y Agrimensura.
- Verástegui 2005. *Seguridad Ciudadana*. Perú: Centro de Investigación Parlamentaria.

## IX. ANEXOS

### 1. Anexo 01: Comparación Técnicas de Minería de Datos vs. Técnicas de Reconocimientos de Patrones

MINERÍA DE DATOS			RECONOCIMIENTO DE PATRONES	
Agrupamiento de datos	Clasificación	Reglas de asociación	K-vecinos más próximos	Redes neuronales
<ul style="list-style-type: none"> <li>• Usado para identificar distintos segmentos de una población y para explicar características comunes de un clúster.</li> <li>• El procedimiento es iterativo, permite la reasignar los objetos.</li> <li>• Agrupa los objetos dentro de clases en base a características, además usa algoritmos de agrupamiento para descubrir una clase de datos previamente desconocida.</li> <li>• Mayor exactitud, ya que el clasificador generaliza mejor.</li> <li>• Resultados más simples de entender.</li> <li>• Se realiza en un solo paso, por lo puede ser muy rápido.</li> </ul>	<ul style="list-style-type: none"> <li>• Basados en aprendizaje supervisado (se tiene información de los grupos o clases a priori).</li> <li>• La asignación de objetos es definitiva.</li> <li>• Operan con una matriz de similitudes.</li> <li>• Categoriza los datos en base a un conjunto de entrenamiento, y hace un modelo para cada clase.</li> <li>• Utilizado para clasificar los nuevos datos agregados a una base de datos.</li> </ul>	<ul style="list-style-type: none"> <li>• Usada para establecer las posibles relaciones entre distintas acciones aparentemente independientes.</li> <li>• Utilizada cuando el objetivo es realizar análisis exploratorios.</li> <li>• Realiza la búsqueda de todos los detalles para encontrar patrones con una alta probabilidad de repetición.</li> <li>• Para su utilización es necesario disponer de información de cada uno de los sucesos llevados a cabo por un mismo individuo.</li> <li>• Aprende con la evolución del sistema, sin necesidad de reprogramación.</li> </ul>	<ul style="list-style-type: none"> <li>• Las reglas de clasificación por vecindad está basada en la búsqueda de un conjunto de modelos de los k modelos más cercanos al patrón a clasificar.</li> <li>• No necesita hacer ninguna suposición sobre los conceptos a aprender.</li> <li>• Presenta versatilidad para manejar patrones con múltiples valores perdidos.</li> <li>• Busca los patrones más cercanos en todo el conjunto de datos, siendo una limitación al ser crítica en grandes bases de datos.</li> </ul>	<ul style="list-style-type: none"> <li>• Necesita pasar por una etapa de entrenamiento para su correcto funcionamiento.</li> <li>• Capaz de predecir con precisión diferentes clases.</li> <li>• La inicialización de los pesos antes del entrenamiento, es delicada ya que una elección inadecuada de valores puede causar un aprendizaje incorrecto.</li> <li>• Alta tolerancia al ruido y fallas en los datos</li> <li>• Requieren de largos tiempos de entrenamiento.</li> <li>• De rápida implantación</li> </ul>

Fuente: (Elaboración Propia 2015).

## 2. Anexo 02: Comparación Técnicas de Supervisadas y No Supervisadas

CLASIFICACIÓN SUPERVISADA			CLASIFICACIÓN NO SUPERVISADA	
Árboles de Decisión	K-vecinos más próximos	Redes Neuronales	K-MEANS	Asociación – A Priori
<ul style="list-style-type: none"> <li>• Su objetivo es obtener reglas que permitan clasificar a partir de atributos.</li> <li>• Las clases deben estar predefinidas.</li> <li>• Pequeño cambio en los datos y puede generar gran cambio en las particiones, generando interpretación precaria.</li> <li>• Recomendable cuando el número de acciones es pequeño y no son posibles todas las combinaciones.</li> <li>• Facilita la interpretación de la decisión adoptada.</li> <li>• Dificultad para elegir un árbol óptimo.</li> <li>• Pueden alcanzar un tamaño considerable, lo cual puede</li> </ul>	<ul style="list-style-type: none"> <li>• Las reglas de clasificación por vecindad está basada en la búsqueda de un conjunto de modelos de los k modelos más cercanos al patrón a clasificar.</li> <li>• No necesita hacer ninguna suposición sobre los conceptos a aprender.</li> <li>• Presenta versatilidad para manejar patrones con múltiples valores perdidos.</li> <li>• Busca los patrones más cercanos en todo el conjunto de datos, siendo una limitación al ser crítica en grandes bases de datos.</li> </ul>	<ul style="list-style-type: none"> <li>• Necesita pasar por una etapa de entrenamiento para su correcto funcionamiento.</li> <li>• Capaz de predecir con precisión diferentes clases.</li> <li>• La inicialización de los pesos antes del entrenamiento, es delicada ya que una elección inadecuada de valores puede causar un aprendizaje incorrecto.</li> <li>• Alta tolerancia al ruido y fallas en los datos</li> <li>• Requieren de largos tiempos de entrenamiento.</li> <li>• De rápida</li> </ul>	<ul style="list-style-type: none"> <li>• Más conocido y utilizado por su simple aplicación y eficacia.</li> <li>• Técnica multivariante, permite agrupar las variables de un archivo de datos en función a la similitud entre ellas.</li> <li>• Permite a un individuo ser reasignado de un grupo a otro, si ello optimiza el criterio de selección.</li> <li>• Permite procesar un número ilimitado de casos, sin embargo requiere que se proponga el número de grupos a obtener.</li> <li>• Tiene capacidad de manejar diferentes tipos de atributos: numéricos, nominales, binarios, ordinales, etc.</li> <li>• Los resultados son</li> </ul>	<ul style="list-style-type: none"> <li>• En los problemas de grandes dimensiones, se entrena más rápidamente.</li> <li>• Facilidad de interpretación de resultados.</li> <li>• Ofrece cinco métodos de entrenamiento distintos, permite mayor flexibilidad para asociarlo con el problema en cuestión.</li> <li>• No analiza patrones, basadas totalmente en la confianza.</li> <li>• El proceso de evaluación de resultados puede ser extenuante.</li> </ul>

complicar el análisis.		implantación	fáciles de entender.	
------------------------	--	--------------	----------------------	--

Fuente: (Elaboración Propia 2015).

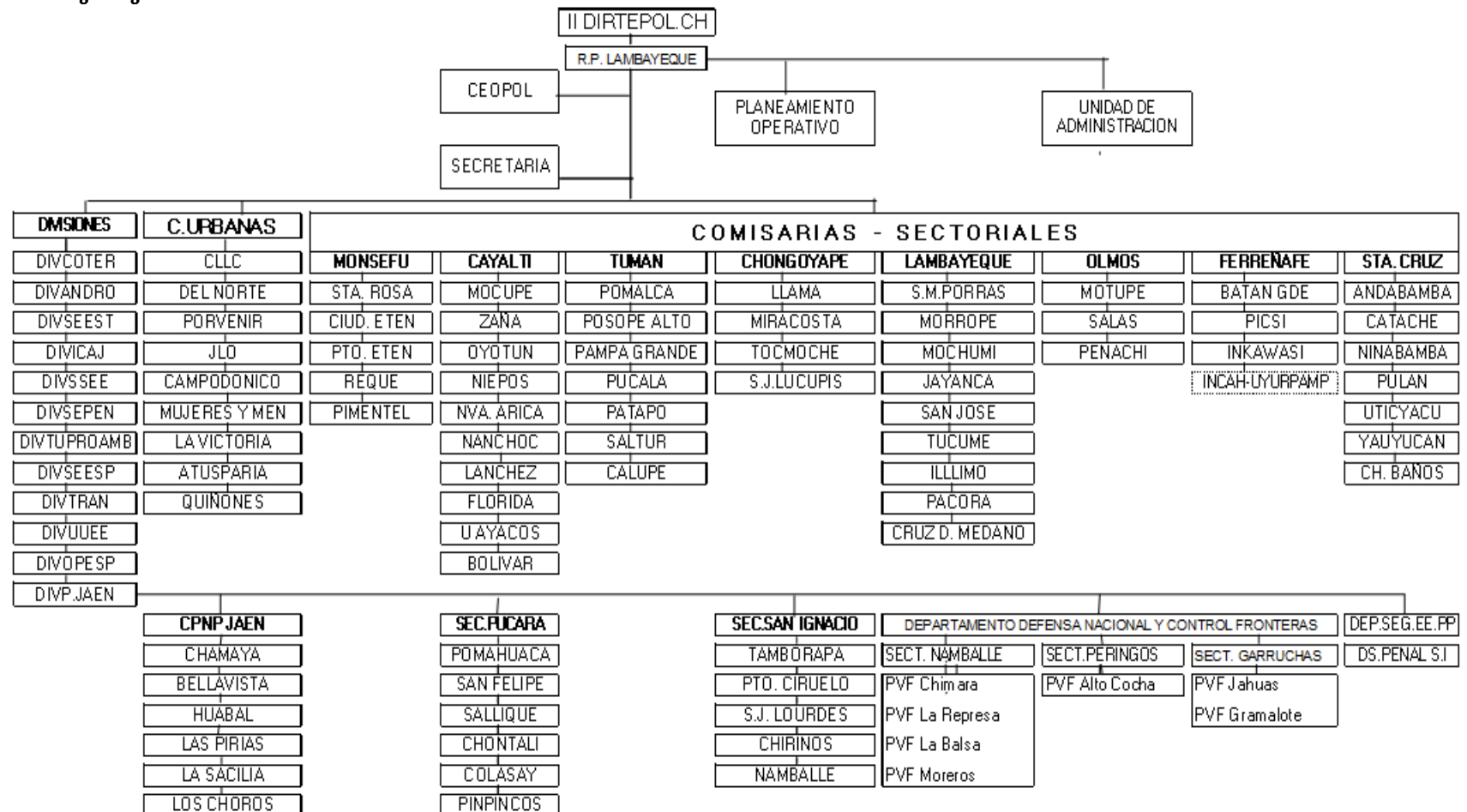
### **3. Anexo 03: Información General sobre la Institución Policial**

#### **a. Descripción**

Región Policial Lambayeque se encuentra ubicada en Balta # 080 - Chiclayo, encargada de realizar estudios e investigaciones estadísticas en el ámbito delictivo, además llevar el control estadístico de las denuncias registradas por todas las comisarías de la ciudad de Chiclayo.

Finalmente tiene como función diseñar programas específicos destinados a actuar sobre la problemática delictiva, tanto preventiva como correctiva, determinando áreas de mayor prioridad, medios idóneos para la producción de las transformaciones que se considere necesarias con relación al tema.

b. Organigrama



4. Anexo 04: Entrevista al SOS PNP Juan Díaz Villegas respecto al Proceso de Registro de Denuncias y Definición de Perfiles Delictivos

Fecha: \_\_\_\_ / \_\_\_\_ / \_\_\_\_

Entrevista N°01

SOS PNP:

\_\_\_\_\_

Pregunta N°01: ¿Cómo se lleva a cabo el proceso de registro de denuncia?

\_\_\_\_\_  
\_\_\_\_\_

Pregunta N°02: ¿Cómo son agrupadas o clasificadas estas denuncias?

\_\_\_\_\_  
\_\_\_\_\_

Pregunta N°03: ¿Qué Medios y Materiales son usados para distribuir el registro y control de las denuncias?

\_\_\_\_\_  
\_\_\_\_\_

Pregunta N°04: ¿Cuáles son los delitos de mayor recurrencia?

\_\_\_\_\_  
\_\_\_\_\_

Pregunta N°05: ¿Se lleva un control de los implicados en un delito de acuerdo a la cantidad de denuncias que tiene?

\_\_\_\_\_  
\_\_\_\_\_

Pregunta N°06: ¿Qué variables o características del implicado se toman en cuenta para determinar un perfil delictivo?

\_\_\_\_\_  
\_\_\_\_\_

Pregunta N°07: ¿Qué Técnicas son utilizadas para evaluar los perfiles delictivos?

\_\_\_\_\_  
\_\_\_\_\_

Pregunta N°08: ¿Cuál es el Sistema de Evaluación usado para determinar los perfiles delictivos?

\_\_\_\_\_  
\_\_\_\_\_

Pregunta N°09: ¿Por qué cree o cuáles son los motivos por el cual la ocurrencia de un hecho delictivo va en aumento en lugar de disminuir?

---



---

Pregunta N°10: ¿Qué medidas de mitigación son utilizadas para combatir los problemas delictivos?

---



---

#### 5. Anexo 05: Encuesta de Satisfacción a los Policías

#### ENCUESTA DE SATISFACCIÓN PARA EL USUARIO DEL SISTEMA DE ALERTA TEMPRANA

Nombre: \_\_\_\_\_

Instrucciones: Lea detenidamente cada ítem y marque con un aspa (X).

ITEM	DESCRIPCIÓN	SI	NO
1.	La aplicación se ejecuta lentamente.		
2.	Recomendaría su utilización dentro de la Institución donde labora		
3.	La aplicación se ha detenido inesperadamente en algún momento		
4.	La aplicación es complicada de usar		
5.	A veces en algunos puntos, no sé cómo continuar		
6.	Disfruto su manejo		
7.	La aplicación cuenta con niveles de acceso		
8.	Existen mensajes que emite el sistema en caso de ingresar con un usuario y/o contraseña inválida		
9.	Se le tienen prohibidas algunas opciones a la aplicación debido al nivel de acceso destinado para usted		
10.	Si el software termina es difícil reiniciarlo.		
11.	Aprender a usarlo me tomó mucho tiempo.		
12.	La manera en que presenta la información es clara y entendible.		
13.	Solo me siento seguro usando comandos y operaciones que me son familiares.		



14.	Nunca tengo la información suficientemente necesaria en pantalla		
15.	Me gustaría usarlo diariamente		
16.	Me brinda ayuda con cualquier problema que me surja		
17.	La organización de los menús o la información listada es poco lógica		
18.	Aprender a usar las nuevas funcionalidades es difícil		
19.	Los mensajes para prevenir errores no son adecuados.		
20.	Tiene una interfaz muy amigable		